DOCUMENT RESUME

ED 338 714 TM 017 599

TITLE	Preliminary Technical Manual for the Enhanced ACT Assessment.
INSTITUTION	American Coll. Testing Program, Iowa City, Iowa.
PUB DATE	Oct 89
NOTE	63p.
AVAILABLE FROM	American College Testing Program Publications, P.O. Box 168, Iowa City, IA 52243.
PUB TYPE	Guides - Non-Classroom Use (055) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*College Entrance Examinations; *Educational Assessment; Educational Change; Equated Scores; Higher Education; High Schools; High School Students; *Interest Inventories; *National Programs; Scaling; *Scoring; Test Construction; *Testing Programs; Test Interpretation; Test Manuals; Test Norms; Test Reliability; Test Validity
IDENTIFIERS	ACT Assessment; ACT Interest Inventory; *Enhanced ACT

ABSTRACT

This manual contains preliminary information, largely of a technical nature, about the enhanced American College Testing (ACT) Program introduced in October 1989. It focuses primarily on the tests of educational development, the component of the program that has been substantially revised. Although one-third of the ACT Interest Inventory Items has been replaced or revised, the score scale and the interpretation procedures remain unchanged. The manual describes: (1) the enhanced ACT Assessment Program and related services; (2) development procedures for the tests of educational development; (3) norming, scaling, and equating the tests of educational development; and (4) the reliability and validity of the tests of educational development. Twenty-two tables and one figure supplement the discussion. Seventeen references are listed. Appendices provide the ACT Interest Inventory and three concordance tables for composite, English, and mathematics scores. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

PATRICIA A. FARRANT

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED338714

Preliminary Technical Manual

for the Enhanced ACT ASSESSMENT

October 1989

ACT

TM017599

BEST COPY AVAILABLE

Preliminary Technical Manual

for the Enhanced ACT ASSESSMENT

October 1989

ACT

ACT endorses the Code of Fair Testing Practices in Education, a statement of the obligations to test takers of those who develop, administer, and use educational tests and data. The Code sets forth criteria for fairness in four areas: developing and selecting appropriate tests, interpreting test scores, striving for fairness, and informing test takers. ACT is committed to ensuring that each of its testing programs upholds the Code's standards for appropriate test development practice and use.

A free copy of the Code may be obtained from ACT Publications, P.O. Box 168, Iowa City, Iowa 52243, 319-337-1429.

©1989 by The American College Testing Program. All rights reserved.

Contents

Contents....ii
Figures....iv
Tables....iv
Preface....v

Chapter 1 The Enhanced ACT Assessment Program and Related Services....1

Overview and Purposes of the Enhanced ACT Assessment Program....1
Code of Fair Testing Practices in Education....1
Background of the ACT Assessment Program....2
Changes in the ACT Assessment....2
Philosophical Basis for the Tests of Educational Development....3
Description of the Assessment Components....4
The Tests of Educational Development....5
The English Test....5
The Mathematics Test....5
The Reading Test....5
The Science Reasoning Test....6
The High School Course/Grade Information Section....6
The Student Profile Section....7
The ACT Interest Inventory....7
Test Dates and Locations....8
Scoring Procedures....8
Procedures for Reporting Scores....9
The Student Report....10
The High School Report....10
The College Report....10
The Score Label....11
The High School List Report....11
Related Research and Information Services....11
The High School Profile Service....12
ACT Research Services....12
The Freshman Class Profile Service....13
The Basic Research Service....13
The Standard Research Service....13

Chapter 2 Development Procedures for the Tests of Educational Development....15

Preparing Test Specifications....15
Content Specifications....15
Statistical Specifications....20
Selection of Item Writers....20
Item Construction....21
Review of Items....21
Item Tryouts....21
Item Analysis of Tryout Units....22
Assembly of New Forms....22
Review Following National Administration....23

Chapter 3 Norming, Scaling, and Equating the Tests of Educational Development.....24

Norms and Score Scale.....24
Sampling.....24
Sample Design and Data Collection.....24
Data Editing.....25
Weighting.....25
Response Rates.....25
Sample Representativeness.....26
Obtained Precision.....27
Scaling.....27
The Score Scale.....28
Scaling Process.....29
Norming.....31
National Study Norms.....31
Additional Norms Reported for 1989-90.....35
Equating.....35

Chapter 4 Reliability and Validity of the Tests of Educational Development.....38

Reliability, Measurement Error, and Effective Weights.....38
Validity.....42
Content Validity.....43
Test Correlations.....43
Comparisons of Groups Differing in Educational Background.....46
Conclusion.....48

References.....49

Appendix A.....51

Appendix B.....53

Figures

4.1 Conditional Standard Error of Measurement for Enhanced ACT Assessment Tests....39

Tables

2.1 Specifications for the ACT English Test....16
2.2 Specifications for the ACT Mathematics Test....17
2.3 Specifications for the ACT Reading Test....18
2.4 Specifications for the ACT Science Reasoning Test....19

3.1 Examinee Sample Sizes....25
3.2 Within-School Student Participation Rates in Proportions....26
3.3 Demographic and Educational Characteristics for Sample 1 of the Academic Skills Study....27
3.4 Properties of the Score Scale for the Enhanced ACT Assessment....29
3.5 Scale Score Summary Statistics for Enhanced ACT Assessment Tests....30
3.6 Scale Score Summary Statistics for Subscores of Enhanced ACT Assessment....31
3.7 Estimated Cumulative Percents on Enhanced ACT Assessment Tests for Fall College-Bound Twelfth-Grade Students....32
3.8 Estimated Cumulative Percents on Enhanced ACT Assessment Tests for Fall National Twelfth-Grade Students....33
3.9 Estimated Cumulative Percents for Enhanced ACT Assessment Subscores for Fall College-Bound Twelfth-Grade Students....34
3.10 Estimated Cumulative Percents for Enhanced ACT Assessment Subscores for Fall National Twelfth-Grade Students....34
3.11 National Percentile Distributions of ACT Assessment Scale Scores, 1986-87-88 High School Graduating Classes....36

4.1 Estimated Reliabilities and Standard Errors of Measurement for Enhanced ACT Assessment Tests....40
4.2 Estimated Reliabilities and Standard Errors of Measurement for Enhanced ACT Assessment Subscores....41
4.3 Scale Score Covariances and Effective Weights....42
4.4 Observed Correlations (Above Diagonal) and Disattenuated Correlations (Below Diagonal) Between Enhanced ACT Assessment Test Scores....44
4.5 Observed Correlations (Above Diagonal) and Disattenuated Correlations (Below Diagonal) Between Enhanced ACT Assessment Subscores....45
4.6 Scale Score Means and Standard Deviations....47
4.7 Means and Standard Deviations for Enhanced ACT Assessment Scores and Subscores: Grade 12 National Groups by Years of Coursework They Have Taken and/or Plan to Take....48

Preface

This manual contains preliminary information, largely of a technical nature, about the enhanced ACT Assessment Program introduced in October 1989. It focuses primarily on the tests of educational development, the component of the program that has been substantially revised. Although one-third of the ACT Interest Inventory items have been replaced or revised, the score scale and the interpretation procedures remain unchanged. Technical information concerning the ACT Interest Inventory, the Student Profile Section, and the High School Course/Grade Information Section can be found in the ACT Assessment Program Technical Manual (American College Testing Program [ACT], 1988).[1] A more extensive and complete technical manual will be available in the future.

[1] All ACT publications cited in this manual are available from the ACT Publications Department, P.O. Box 168, Iowa City, IA 52243, 319-337-1429.

CHAPTER 1

THE ENHANCED ACT ASSESSMENT PROGRAM AND RELATED SERVICES

Overview and Purposes of the Enhanced ACT Assessment Program

The enhanced ACT Assessment Program, like the original ACT Assessment Program, is a comprehensive system of data collection, processing, and reporting designed to help students develop postsecondary plans and to help postsecondary educational institutions develop programs suited to the needs and characteristics of their applicants. The program helps students identify and develop realistic plans for accomplishing their educational and career goals as they move from secondary to postsecondary education. The enhanced ACT Assessment Program information presents students to postsecondary institutions as persons with unique patterns of educational development, accomplishments, and needs. High schools use ACT Assessment data in academic advising and counseling. Colleges use ACT Assessment results in the admissions process and in placing students into instructional programs. Many of the agencies that provide scholarships, loans, and other types of financial assistance to students tie such assistance to students' academic qualifications. And many state agencies conduct recognition programs for state scholars. ACT Assessment data are provided for all these purposes. Recently, various state and federal agencies and some educational consortia have come to rely on ACT Assessment data as elements in their overall educational evaluation and planning activities.

The enhanced ACT Assessment Program consists of (1) a battery of four tests of educational development; (2) a questionnaire relating to high school courses and grades; (3) a questionnaire on students' educational and career aspirations, extracurricular activities, and special educational needs; and (4) an interest inventory. The tests are taken under standardized conditions; the other components are completed when the students register to take the tests.

Code of Fair Testing Practices in Education

In 1988, ACT endorsed the Code of Fair Testing Practices in Education, a statement of the obligations to test takers of those who develop, administer, and use educational tests and test data. The development of the Code was sponsored by a joint committee of the American Association for Counseling and Development, Association for Measurement and Evaluation in Counseling and Development, American Educational Research Association, American Psychological Assocation, American Speech-Language-Hearing Association, and National Council on Measurement in Education to advance, in the public interest, the quality of testing practices.

The Code sets forth fairness criteria in four areas: developing and selecting appropriate tests; interpreting scores; striving for fairness to test takers of different races, sexes, ethnic backgrounds, or handicapping conditions; and providing information to test takers about tests and about test takers' rights. Separate standards are provided for test developers and for test users in each of these four areas.

ACT's endorsement of the Code represents a commitment to vigorously safeguard the rights of individuals participating in its testing programs. ACT has designed an ongoing review process whereby each of its testing programs is routinely reviewed to ensure that it upholds the standards set forth in the Code for appropriate test development practice and test use.

A copy of the Code may be obtained free of charge from ACT Publications, P.O. Box 168, Iowa City, Iowa 52243, 319-337-1429.

Background of the ACT Assessment Program

The original ACT Assessment Program was designed in the late 1950s to serve the needs of large state universities; state, municipal, and junior colleges; and the great majority of the smaller private and denominational colleges of the country. Growth in the college-bound student population in the early 1960s made it necessary that most colleges acquire meaningful and objective data on which to base decisions about students. Institutions with open admissions policies valued information useful in counseling students about programs and course placement. Schools with selective admissions policies needed information useful in identifying those students most likely to benefit from enrollment in their educational programs. Thus, the institutions needed a test battery and information services that provided a comprehensive description of their students' educational needs and achievement level.

Because the goals of the ACT Assessment are closely tied to the goals and practices of high schools and colleges, ACT has over the past three decades been especially sensitive to trends and expectations at these levels, and has regularly modified the ACT Assessment Program to ensure its continued sensitivity to changing needs in secondary and postsecondary education.

The enhanced ACT Assessment, introduced in October 1989, is a revised program that is responsive to changes that have occurred in high school curricula, is sensitive to current expectations about the skills and knowledge students need for success in college, and offers information educators need to address some of the many challenges they now face.

To make the changes in the ACT Assessment as powerful and useful as possible, ACT sought the advice and opinions of college and high school teachers and administrators, course content experts, and curriculum specialists; scrutinized various critiques of American education; and reviewed state curriculum requirements and widely used high school and college textbooks. The enhanced ACT Assessment reflects this careful, methodical study and speaks to many of the needs it revealed.

Changes in the ACT Assessment

Some of the major findings the above process produced include the following:

- The need for improved information for placement and advising in English, mathematics, reading, and science

- The need for a measure of "pure" reading ability and comprehension
- The importance of scientific literacy and reasoning skills
- The limited use of the current ACT Social Studies Reading and Natural Sciences Reading scores

The content of the four ACT tests of educational development has been revised to address these concerns. The enhanced ACT Assessment increases the emphasis on rhetorical skills in the measurement of writing proficiency, increases the number of advanced math items, and includes both a new reading test which features inferential and reasoning skills, and a test designed to measure science reasoning. The enhanced ACT Assessment will also provide subscores in English, Mathematics, and Reading--more detailed information about a student's level of educational development.

All other components of the ACT Assessment remain essentially unchanged. Although more than one-third of the items in the ACT Interest Inventory have been replaced or revised, the score scale and interpretation procedures for the inventory have not been changed.

Philosophical Basis for the Tests of Educational Development

Underlying the ACT Assessment tests of educational development is the philosophy that student preparedness for college is best assessed by measuring, as directly as possible, the academic skills that the student will need to perform college-level work. The required academic skills can be most directly assessed by reproducing as faithfully as possible the complexity of college-level work. Therefore, the tests of educational development are designed to determine how skillful the student is in solving problems, grasping implied meanings, drawing inferences, evaluating ideas, and making judgments.

In addition, the tests of educational development are oriented around the general content areas of college and high school instructional programs. The test questions require students to integrate the knowledge and skills they possess in major curriculum areas with the information provided by the test. In this way, scores on the tests have a direct and obvious relationship to the students' educational progress in curriculum-related areas and possess a meaning that is readily grasped by students, parents, and educators.

Tests of general educational development are used on the ACT Assessment because, when compared to other types of tests, they best satisfy the diverse requirements of tests used to facilitate the transition from secondary to postsecondary education. By contrast, measures of examinee knowledge of specific course content (as opposed to curriculum areas) do not readily provide a common baseline for comparing students for the purposes of admission, selection, or awarding scholarships because high school courses vary so extensively. In addition, such tests might not measure students' skills in problem solving and in the integration of knowledge from a variety of courses.

Tests of educational development can also be contrasted with tests of academic aptitude. The stimuli and test questions for aptitude tests are

often chosen precisely for their dissimilarity to instructional materials, and each test within a battery of aptitude tests is designed to be homogeneous in psychological structure. With such an approach, these tests might not reflect the complexity of college-level work or the interactions among the skills measured. In addition, because aptitude tests are not directly related to instruction, they might not be as useful as tests of educational development for making placement decisions in college.

The advantage of tests of educational development over other types of tests for use in the transition from high school to college becomes evident when their use is considered in the context of the educational system. Because many of the same skills that are taught in high school are being measured, the best preparation for tests of educational development is high school coursework. Long-term learning in school, rather than short-term cramming and coaching, becomes the best form of test preparation. Thus, tests of educational development tend to serve as motivators by sending students a clear message that high test scores are not simply a matter of innate ability but reflect a level of achievement that has been earned as a result of hard work and dedication.

Because the ACT Assessment stresses such general concerns as the complexity of college-level work and the integration of knowledge from a variety of sources, students may be influenced to acquire skills necessary to deal with these concerns. In this way, the ACT Assessment may serve to aid high schools in developing in their students the higher-order thinking skills that are important for success in college and later life.

The tests of the ACT Assessment are designed to accurately reflect educational goals that are widely accepted and judged by educators to be important for success in college. As such, the content of the ACT Assessment is driven by educational considerations, with statistical and empirical techniques being of secondary importance.

Description of the Assessment Components

Four components comprise the data collection part of the enhanced ACT Assessment Program: (1) a battery of four tests of educational development; (2) a questionnaire relating to high school courses and grades; (3) a questionnaire on students' educational and career aspirations, extracurricular activities, and special educational needs; and (4) an interest inventory. The test battery yields scores and subscores in four academic areas (English, mathematics, reading, and science reasoning) and a Composite score. The High School Course/Grade Information Section of the enhanced ACT Assessment Program collects students' self-reported high school grades. The Student Profile Section contains questions relating to students' educational and vocational plans and needs, extracurricular interests and accomplishments, and background characteristics. An analysis of student preferences for six career areas (science, arts, social service, business contact, business operations, and technical) results from a student's responses to items on the ACT Interest Inventory.

Students who take the tests of educational development on one of the national test dates complete the three noncognitive components as part of the

registration procedure. Other (e.g., residually tested) students may or may not complete all three of these sections.

The Tests of Educational Development

The English Test is a 75-item, 45-minute test that measures the students' understanding of the conventions of standard written English (punctuation, grammar and usage, and sentence structure) and of rhetorical skills (strategy, organization, and style). Spelling, vocabulary, and rote recall of rules of grammar are not tested. The test consists of five prose passages, each of which is accompanied by a sequence of multiple-choice test items. Different passage types are employed to provide a variety of rhetorical situations. Passages are chosen not only for their appropriateness in assessing writing skills, but also to reflect the interests and experiences of examinees. Some items refer to underlined portions of the text and offer several alternatives to the portions underlined. Other items pose questions about a section of the text, or about the passage as a whole. The student must decide which alternative is most appropriate in the context of the passage, or which alternative best answers the question posed. Many items offer as one alternative response "NO CHANGE" from the text.

Three scores are reported for the ACT English Test: a total test score based on all 75 items, a subscore in Usage/Mechanics based on 40 items, and a subscore in Rhetorical Skills based on 35 items.

The Mathematics Test is a 60-item, 60-minute test that is designed to assess the mathematical skills that students have typically acquired in courses taken up to the beginning of grade 12. The test presents multiple-choice items that require students to use their reasoning skills to solve practical problems in mathematics. Knowledge of basic formulas and computational skills are assumed as background for the problems, but complex formulas and extensive computation are not required. The material covered on the test emphasizes the major content areas that are prerequisite to successful performance in entry-level courses in college mathematics.

The items included in the Mathematics Test cover three skill areas: (1) basic skills, (2) application, and (3) analysis. Basic skills includes items that can be solved by performing a familiar sequence of operations in a familiar setting. Application items can be solved by performing a familiar sequence of operations, but the solution will not be routine. Analysis items require a student to know why the familiar sequence of operations yields a solution, under what conditions it will not yield a solution, or how to examine all the cases that can arise within the restrictions stated in the stem of the item.

Four scores are reported for the ACT Mathematics Test: a total test score based on all 60 items, a subscore in Pre-Algebra/Elementary Algebra based on 24 items, a subscore in Intermediate Algebra/Coordinate Geometry based on 18 items, and a subscore in Plane Geometry/Trigonometry based on 18 items.

The Reading Test is a 40-item, 35-minute test that measures reading comprehension as a product of skill in referring and reasoning. That is, the test items require students to derive meaning from several texts by

(1) referring to what is explicitly stated and (2) reasoning to determine implicit meanings and to draw conclusions, comparisons, and generalizations. The test comprises four prose passages (based on topics in prose fiction, the humanities, the social studies, and the natural sciences) that are representative of the level and kinds of writing commonly encountered in college freshman curricula. These passages are selected from published sources. Each passage is accompanied by a set of multiple-choice test items. These items do not test the rote recall of facts from outside the passage, isolated vocabulary items, or rules of formal logic. Rather, the test focuses upon the complex of complementary and mutually supportive skills that readers must bring to bear in studying written materials across a range of subject areas.

Three scores are reported for the ACT Reading Test: a total test score based on all 40 items, a subscore in Social Studies/Science reading skills (based on the 20 items in the social studies and natural sciences sections of the test) and a subscore in Arts/Literature reading skills (based on the 20 items in the prose fiction and humanities sections of the test).

The Science Reasoning Test is a 40-item, 35-minute test that measures the interpretation, analysis, evaluation, reasoning, and problem-solving skills required in the natural sciences. The test is made up of test units, each of which consists of some scientific information and a set of multiple-choice test items. The scientific information is conveyed in one of three different formats: data representation (graphs, tables, and other schematic forms), research summaries (descriptions of several related experiments), or conflicting viewpoints (expressions of several related hypotheses or views that are inconsistent with one another).

The test items require students to recognize and understand the basic features of and concepts related to the provided information; to examine critically the relationships between the information provided and the conclusions drawn or hypotheses developed; and to generalize from given information to gain new information, draw conclusions, or make predictions.

The content of the Science Reasoning Test is drawn from Biology, Chemistry, Physics, and the Physical Sciences (e.g., Geology, Astronomy, and Meteorology). Advanced knowledge in these subjects is not required, but background knowledge at the level of a high school General Science course may be needed to answer some of the questions. Advanced mathematical skills are not required, but minimal arithmetic computations may be needed for some questions. Thus, the test emphasizes scientific reasoning skills rather than recall of scientific content, skill in mathematics, or reading ability.

A total test score only is reported for the ACT Science Reasoning Test.

The High School Course/Grade Information Section

Information about courses taken and grades received is collected on this section of the ACT Assessment Registration Folder. Students are asked to indicate whether they have taken or plan to take each of 30 different courses. They are also asked to record the last grade earned in each completed full-term course. In order to minimize errors in these self-reported data, the instructions advise students to refer to grade reports or

transcripts. They are also asked to sign a statement attesting to the accuracy of the reported grades. Information concerning the accuracy of self-reported high school courses and grades can be found in Sawyer, Laing, and Houston (1988). A copy of the instructions for completing this section of the ACT Assessment Registration Folder are provided in the ACT Assessment Program Technical Manual (American College Testing Program [ACT], 1988). Information concerning the development and validity of the High School Course/Grade Information Section can be found in the same publication.

The Student Profile Section

The Student Profile Section (SPS) of the ACT Assessment Registration Folder asks the students to supply information about their backgrounds, interests, needs, and plans. It is designed to help students think about their educational future and to help postsecondary institutions match the needs of students with the programs they offer. Students are informed that information about racial/ethnic background, native language, marital status, religious preference, and physical disability is released only with their consent to institutions that request it in accordance with federal guidelines, and that they are not required to provide this information.

The SPS also asks for nearly all of the information colleges typically request on their application forms. The student responds to questions about enrollment plans; proposed educational major; level of educational degree sought; college housing needs; vocational plans; need for assistance in improving skills in studying, writing, reading, or mathematics; and interest in credit by examination and advanced placement. Some institutions use the ACT Assessment record as a total admission system (ACT, 1982), thus eliminating the need for students to submit separate admission applications and/or high school transcripts.

Also collected on the SPS are the following: demographic data; high school information, including number of years specific subjects were studied in high school; and notable high school accomplishments in the areas of leadership, music, speech, art, writing, science, athletics, work experience, and community service.

Although students complete the SPS on a voluntary basis, most recognize the importance of providing such information to colleges and do complete it. A response rate greater than 95% was recently reported (ACT, 1987). The questions included in the SPS and instructions for answering them are provided in the ACT Assessment Program Technical Manual (ACT, 1988). Also provided in the same publication is information concerning the development procedures, scoring procedures, norms, validity evidence, and reliability data of the SPS.

The ACT Interest Inventory

The ACT Interest Inventory is a 90-item instrument designed to measure students' work task preferences in six comprehensive career areas: Science, Arts, Social Service, Business Contact, Business Operations, and Technical. Each of the six scales is composed of 15 work-related activities for which students indicate their degree of liking on a three-point scale (like, indifferent, dislike). The ACT Interest Inventory scales were developed to parallel Holland's (1985) six interest and occupational types.

Several different but related interest inventories have been developed for and used in various ACT programs. The current form is the Unisex Edition of the ACT Interest Inventory (UNIACT). The UNIACT was constructed with the goal that the same distribution of career options would be suggested to men and women. UNIACT was introduced in 1977 and has been updated for the fall of 1989. More than one-third of the items in UNIACT have been replaced or revised. Item replacements and revision were made on the basis of item tryouts involving 2,000 ninth graders, 5,000 eleventh graders (including students taking the ACT Assessment), 2,000 twelfth graders, and 3,000 adults. The score scale and the interpretation procedures remain unchanged and can be found in the ACT Assessment Program Technical Manual (ACT, 1988). The items on the inventory and directions for responding to them are provided in Appendix A of this manual.

Test Dates and Locations

Each year the enhanced ACT Assessment Program is administered throughout the United States on five Saturday test dates. There are also regularly scheduled non-Saturday test dates for students whose religious faith prohibits them from taking a test on Saturday. Special administrations are arranged for those confined to hospitals or in prisons, or students with diagnosed physical or learning disabilities. Special administrations can also be arranged for those living more than 50 miles from all non-Saturday test centers on those scheduled test dates. The enhanced ACT Assessment Program is also administered on four test dates each year in countries all over the world. Personnel on active duty at United States military bases take the enhanced ACT Assessment through DANTES (Defense Activity for Nontraditional Education Support). National test dates and locations (including non-Saturday locations) are published annually in Registering for the ACT Assessment. ACT strives to provide enough centers to test all students who wish to participate in the enhanced ACT Assessment Program. Test dates, registration procedures, fees, and deadlines for test centers in other countries are published in Taking the ACT Assessment for Students Outside the 50 United States.

There is also an on-campus testing program for ACT Assessment-participating postsecondary institutions that have admitted or enrolled students who did not take the enhanced ACT Assessment tests on one of the national test dates. The results of this type of testing, referred to as residual testing, are reported *only* to the institution administering the tests.

Scoring Procedures

For each of the four content area tests in the enhanced ACT Assessment (English, Mathematics, Reading, Science Reasoning), the raw scores (number of correct responses) are converted to scale scores that range from 1 to 36. Raw scores for each of the seven subscores are converted to scale scores on a scale that ranges from 1 to 18. These score scales are discussed further in Chapter 3 of this manual. The Composite score is the average of the four scale scores rounded to the nearest whole number (.5 rounded up). The minimum Composite score is 1, the maximum 36.

Electronic scanning devices are used to score all parts of the enhanced ACT Assessment Program, thus minimizing the potential for scoring errors. If

a student believes that a scoring error has been made, ACT handscores the answer sheet at no extra charge upon receipt of a written request from the student. A student may arrange to be present for handscoring by contacting one of ACT's regional offices but must pay whatever extra costs may be incurred in providing this special service. Strict confidentiality of each student's record is maintained.

For certain test dates (specified in Registering for the ACT Assessment), each examinee may obtain (upon payment of an additional fee) a copy of the test questions used in determining his/her scores, a list of his/her answers, and a table to convert raw scores to the reported scale scores. For an additional fee, a student may also request a copy of his/her answer sheet. These materials are available only to students who test during regular administrations of the enhanced ACT Assessment tests on specified national test dates. If for any reason ACT must replace the test form scheduled for use at a test center, this offer is withdrawn and the student's fee for this optional service is refunded.

ACT reserves the right to cancel test scores when there is reason to believe the scores are invalid. Cases of irregularities in the test administration process--for example, impersonating another examinee (surrogate testing), unusual similarities in the answers of examinees at the same test center, and other indications, including but not limited to student misconduct, that the test scores may not accurately reflect the student's level of educational development--may result in ACT's canceling the test scores. When ACT plans to cancel an examinee's test scores, it always notifies the examinee prior to taking that action. This notification includes information about alternatives to score cancellation, including procedures for appealing that step. In all instances, the final and exclusive remedy available to examinees who want to appeal or otherwise challenge a decision by ACT to cancel their test scores shall be binding arbitration through written submissions to the American Arbitration Association.

Procedures for Reporting Scores

The set of reports described in this section is distributed four to five weeks after the test date. No other copy of a student's report, or any part of it, is released without the student's authorization.

Information from the enhanced ACT Assessment Program is aggregated in various forms based on the intended recipient and then reported to each. Unique reports are provided for the student (Student Report), the high school (High School Report), and to as many as six universities and other agencies (College Report) selected by the student on the registration form. Score Labels and a High School List Report are also sent to the high school that the student indicated he/she attends. These reports are described briefly below and in more detail in the ACT document The ACT Assessment User Handbook. The Student Report is accompanied by Using Your ACT Assessment Results, a booklet designed to help students with their educational and vocational planning.

The Student Report

The Student Report provides the student with an individualized narrative summary of his or her main results. The report presents the four test scores, the Composite score, and the seven subscores along with their corresponding national college-bound ranks expressed as cumulative percents--the percent of examinees in the Fall 1988 Academic Skills Study (see Chapter 3) who scored at or below each score. When it is possible to do so, the report provides additional ranks for the student's Composite score compared to other groups of college-bound examinees. For 1989-90, one of these groups will be ACT Assessment-tested high school graduates of 1986, 1987, and 1988. These norms, discussed on page 35 of this manual, were determined by using enhanced ACT Assessment scores estimated from a concordance table (see Appendix B) linking the original ACT Assessment to the enhanced ACT Assessment. The report also conveys information about up to six college choices and personalized suggestions for educational and career planning.

The High School Report

The High School Report provides counselors with test scores, subscores, and ACT Interest Inventory results plus detailed information about the student's plans, needs, and goals beyond high school. Information about the student's college choices, which appears on both the High School Report and the Student Report, is furnished for some 3,500 colleges that respond to ACT's annual Institutional Data Questionnaire and authorize release of their data.

For 1989-90, the High School Report describes test scores with respect to two reference groups. The percent of college-bound students who earned a scale score at or below that that the examinee earned, based on the Fall 1988 Academic Skills Study (see Chapter 3), is shown for all reported scores. In addition, estimated norms for the English, Mathematics, and Composite scores, based on ACT Assessment-tested high school graduates of 1986, 1987, and 1988, are given. These norms were determined by using enhanced ACT Assessment scores estimated from concordance tables (see Appendix B) linking the original ACT Assessment to the enhanced ACT Assessment.

Students' ACT Assessment records, as shown on the High School Report, are available to high schools on magnetic tape as an extra-cost reporting service. All schools first receive the standard paper copies of the High School Report for use in counseling individual students. Tapes are generated when all scoring and matching has been completed, generally six to eight weeks after each test date. Order forms with specifications and prices are available from ACT Reporting Services.

The College Report

Similar in format and content to the High School Report, the College Report provides admissions, advising, and placement staff with additional information about the likelihood of a student earning a "C" or better in specific courses and/or academic majors.

Based on the Fall 1988 Academic Skills Study (see Chapter 3), the percent of college-bound students who earned a scale score at or below each of the twelve reported scores is listed. In addition, based on local institutional

research, estimated norms are provided for English, Mathematics, and the Composite score. These norms were determined by applying the concordance tables (see Appendix B) to link the previous ACT Assessment scores to enhanced ACT Assessment scores. As a result, for score reports prepared in 1989-90, it was possible to provide estimated enhanced ACT Assessment norms based on the previous year's ACT Assessment-tested freshmen. In addition, a footnoote on the report provides estimated norms for English, Mathematics, and Composite scores based on 1986-1988 ACT Assessment-tested high school graduates. These norms were determined using the concordance tables (see Appendix B).

Students' ACT Assessment records, as shown on the College Report, are available to colleges on magnetic tape. The tape format for 1989-90 has been revised to include the seven subscores. The revised format also includes estimated scores on the enhanced ACT Assessment for students who tested before October 1989 and later request Additional Score Reports. A layout of the revised format and information about ordering test tapes were mailed to colleges in 1988.

Tapes formerly were available only after national test dates and at year-end cleanup. Beginning in October 1989, weekly and monthly tapes are available. Details about these and other reporting options--including a new teletransmission option--are provided in the brochure Reporting Services for Colleges and Universities, which is mailed to colleges in July of each year.

The Score Label

High schools receive two score labels for each student who correctly entered the school's code in completing the ACT Assessment. The labels bear the student's name, Social Security number, test date, test scores, Composite score, and subscores. One label is provided for the student's permanent record; the other, for the counselor's files.

The High School List Report

High schools also receive an alphabetical list report of all students who indicated that school's code. The report lists students alphabetically within the grade levels that they noted on their registration folders.

The report includes students' test scores, Composite score, and subscores, with cumulative percents for the test scores and Composite score based on the ACT national study college-bound norm group. Other data reported include students' chances in 10 of earning a first-term average of "C" or higher at their first three college choices. The college major and first occupational choice are reported as three-digit codes found in Registering for the ACT Assessment, and the student's response to the item (SPS 190) releasing his/her record for inclusion in ACT's Educational Opportunity Service is noted.

Related Research and Information Services

This section contains brief overviews of various services that are provided by ACT in addition to the reports described in the last section. More information about these services may be found in the booklet Research and

Information Services and in the specific manuals for each service referred to in this section.

The High School Profile Service

The ACT High School Profile Service is provided to high schools from which 40 or more students in the most recent ***graduating class*** took the ACT Assessment and listed their correct high school code. For students who have tested more than once, only the most recent testing is included.

Each eligible school receives descriptive group information about its students for use in counseling and in program planning and evaluation. This service includes the ACT High School Profile Report, which presents statistical tables that describe various characteristics of the students tested. The following areas are covered: educational development, goals, and aspirations; student body characteristics; evaluation of the local high school; and college-choice data. The report also contains a five-year summary of major findings and trends.

In September a copy of the ACT High School Profile Report is sent automatically, free of charge, to each eligible school. This report includes state norms for the high school's state (only available in states that test a sufficient number of students), regional norms, and national norms. Also included in this mailing are the following:

- A set of charts to be completed and used in presenting the report data to others
- A copy of College Student Profiles (revised periodically and included only if updated in the current year)
- A copy of Your College-Bound Students for use in interpreting tables in the report (revised periodically and included only if updated in the current year)

Five-year trend data for average estimated enhanced ACT Assessment English, Mathematics, and Composite scores will be developed and distributed to high schools in Fall 1989. Trend data will be provided for 1984-85, 1985-86, 1986-87, 1987-88, and 1988-89 ACT-tested high school graduates.

The High School Profile Report distributed in September 1990 will be based on enhanced ACT Assessment scores. Local, state, regional, and national norms will also be based on enhanced ACT Assessment scores.

ACT Research Services

ACT Research Services are designed to aid colleges in research, self-study, and planning by providing uniform analyses of student data through a variety of descriptive, predictive, and evaluative options. Research Services reports simplify, for participating colleges, the tasks of developing normative data, examining their programs and practices, and comparing their findings with those obtained from other colleges.

Any college that meets eligibility criteria may participate free of charge in the Freshman Class Profile Service and either the Basic Research Service or the Standard Research Service. Descriptions and schedules for these services are mailed to colleges each August in the booklet Research and Information Services. Information also is available from the ACT Research Services Department and ACT regional offices.

The Freshman Class Profile Service. This service provides a comprehensive description of the college's entering freshmen and a parallel description of students who had ACT Assessment reports sent to the college but did not enroll there. Profile descriptions also can be obtained for two specified groups of enrolled freshmen.

Drawing on test scores, high school grades, and other educational information collected by the ACT Assessment, a series of tables describes the following characteristics of members of the freshman groups:

- Academic abilities, goals, and aspirations
- Selected needs for student personnel services
- Out-of-class accomplishments
- College selection considerations
- Demographic characteristics

Other tables provide cross-tabulations among student characteristics and highlight year-to-year trends. The tables are preceded by a narrative summary of major findings.

The Basic Research Service. One of two ACT Assessment prediction services available to colleges, the Basic Research Service provides analyses of the relation between ACT Assessment data (scores and high school GPA) and first-term overall college grades. The analyses are used to develop prediction equations for forecasting the performance of prospective students.

The Basic Research Service minimizes reporting responsibilities for the colleges and provides a brief, easily understood report. Small colleges, newly participating colleges, and colleges with small research staffs often prefer this plan to the Standard Research Service.

The Standard Research Service. This service provides a description of the academic abilities and achievements of the current freshman class and develops prediction equations used in forecasting the performance of future students. Colleges receive a comprehensive analysis of students' precollege information, including ACT Assessment scores, high school grades, and, if desired, locally collected variables. The analysis also includes first-year or first-term achievement for as many as nine groups of students. A summary report accompanies the analysis. Overall grades and grades in four specific courses or subject areas are used.

The Standard Research Service makes possible the analysis of as many as 40 specific course grades in addition to 10 overall GPAs. As in the Basic

Research Service, regression equations developed in the analyses are used to furnish academic predictions on the ACT Assessment Student Report, High School Report, and College Report in subsequent years.

A college may also choose to report as many as five locally collected measures for each student group. Research Service analyses then provide comparisons of these measures with standard ACT predictors.

CHAPTER 2

DEVELOPMENT PROCEDURES FOR THE TESTS OF EDUCATIONAL DEVELOPMENT

In this chapter the procedures used in the development and revision of the tests of educational development in the enhanced ACT Assessment Program are described. The development and revision procedures used on the non-cognitive components of the enhanced ACT Assessment Program are described in the ACT Assessment Program Technical Manual (ACT, 1988).

The test development cycle required to produce each new form of the tests of educational development takes about 2½ years and involves several stages, beginning with preparation of the test specifications.

Preparing Test Specifications

Two types of test specifications are used in the development of the enhanced ACT Assessment tests: content specifications and statistical specifications.

Content Specifications

The enhanced ACT Assessment tests are based on and oriented toward the major areas of secondary and postsecondary instructional programs. The specific knowledge and skills selected for evaluation were determined through a detailed analysis of three sources of information: (1) instructional objectives for grades 7 through 12 for all states in the United States that had published such objectives, (2) textbooks on state-approved lists for courses in grades 7 through 12, and (3) educators at secondary and postsecondary levels who were consulted to determine the knowledge and skills taught in grades 7 through 12 that are prerequisite to successful performance in postsecondary courses.

Based on the knowledge and skills selected through the process described above, content specifications for the enhanced ACT Assessment tests were developed with the assistance of nationally known education consultants. The consultants met with ACT staff to determine both the specific topics to be covered by the tests, and the proportion of items needed to measure these topic areas. The content coverage and the number of items included in each test were selected to reflect the emphasis of the high school and postsecondary curricula.

While care will be taken to ensure that the basic structure of the enhanced ACT Assessment tests will remain the same from year to year so the scores are comparable, the specific allocation of test items to the tests will be "fine-tuned" on a yearly basis. Each year, consultant panels will be convened to review the new forms of the tests in order to verify their content accuracy and the match of the content of the tests to the content emphasis of the secondary curriculum and the requirements of postsecondary coursework. Periodically, national curriculum surveys will also be conducted to identify shifts and changing emphases in the secondary and postsecondary curricula. The basic structure of the content specifications for each of the enhanced ACT Assessment tests is provided in Tables 2.1 through 2.4.

Table 2.1

Specifications for the ACT English Test

Description of the test. The English Test is a 75-item, 45-minute test that measures the students' understanding of the conventions of standard written English (punctuation, grammar and usage, and sentence structure) and of rhetorical skills (strategy, organization, and style). Spelling, vocabulary, and rote recall of rules of grammar are not tested. The test consists of five prose passages, each of which is accompanied by a sequence of multiple-choice test items. Different passage types are employed to provide a variety of rhetorical situations. Passages are chosen not only for their appropriateness in assessing writing skills, but also to reflect the interests and experiences of examinees. Some items refer to underlined portions of the text and offer several alternatives to the portions underlined. Other items pose questions about a section of the text, or about the passage as a whole. The student must decide which alternative is most appropriate in the context of the passage, or which alternative best answers the question posed. Many items offer as one alternative response "NO CHANGE" from the text.

Three scores are reported for the ACT English Test: a total test score based on all 75 items, a subscore in Usage/Mechanics based on 40 items, and a subscore in Rhetorical Skills based on 35 items.

Content of the test. Six elements of effective writing are included in the English Test. Definitions of these elements and the approximate proportion of the test devoted to each are given below.

Content/Skills	Proportion of Test	Number of Items
Usage/Mechanics	.53	40
Punctuation	(.13)	(10)
Grammar and Usage	(.16)	(12)
Sentence Structure	(.24)	(18)
Rhetorical Skills	.47	35
Strategy	(.16)	(12)
Organization	(.15)	(11)
Style	(.16)	(12)
Total	1.00	75

1. **Usage/Mechanics**

   a. **Punctuation.** The items in this category test the conventions of internal and end-of-sentence punctuation, with emphasis on the relationship of punctuation to meaning (e.g., avoiding ambiguity, identifying appositives).

   b. **Basic Grammar and Usage.** The items in this category test agreement between subject and verb, between pronoun and antecedent, and between modifiers and the words modified; verb formation; pronoun case; formation of comparative and superlative adjectives and adverbs; and idiomatic usage (e.g., choosing appropriate function words).

   c. **Sentence Structure.** The items in this category test relationships between and among clauses, placement of modifiers, and shifts in construction.

2. **Rhetorical Skills**

   a. **Strategy.** The items in this category test the appropriateness of expression in relation to audience and purpose, the strengthening of writing with appropriate supporting material, and the effective choice of statements of theme and purpose.

   b. **Organization.** The items in this category test the organization of ideas and the relevancy of statements in context (order, coherence, unity).

   c. **Style.** The items in this category test precision and appropriateness in the choice of words and images, rhetorically effective management of sentence elements, avoidance of ambiguous pronoun references, and economy in writing.

Table 2.2

Specifications for the ACT Mathematics Test

Description of the test. The Mathematics Test is a 60-item, 60-minute test that is designed to assess the mathematical skills that students have typically acquired in courses taken up to the beginning of grade 12. The test presents multiple-choice items that require students to use their reasoning skills to solve practical problems in mathematics. Knowledge of basic formulas and computational skills are assumed as background for the problems, but complex formulas and extensive computation are not required. The material covered on the test emphasizes the major content areas that are prerequisite to successful performance in entry-level courses in college mathematics.

The items included in the Mathematics Test cover three skill areas: (1) basic skills, (2) application, and (3) analysis. Basic skills includes items that can be solved by performing a familiar sequence of operations in a familiar setting. Application items can be solved by performing a familiar sequence of operations, but the solution will not be routine. Analysis items require a student to know why the familiar sequence of operations yields a solution, under what conditions it will not yield a solution, or how to examine all the cases that can arise within the restrictions stated in the stem of the item.

Four scores are reported for the ACT Mathematics Test: a total test score based on all 60 items, a subscore in Pre-Algebra/Elementary Algebra based on 24 items, a subscore in Intermediate Algebra/Coordinate Geometry based on 18 items, and a subscore in Plane Geometry/Trigonometry based on 18 items.

Content of the test. Items are classified according to five content categories. These categories and the approximate proportion of the test devoted to each are given below.

Content Area	Proportion of Test	Number of Items
Pre-Algebra and Elementary Algebra	.40	24
Intermediate Algebra and Coordinate Geometry	.30	18
Plane Geometry	.23	14
Trigonometry	.07	4
Total	1.00	60

a. **Pre-Algebra.** Items in this category are based on operations with whole numbers, decimals, fractions, and integers. They also may require the solution of linear equations in one variable.

b. **Elementary Algebra.** Items in this category are based on operations with algebraic expressions. The most advanced topic in this category is the solution of quadratic equations by factoring.

c. **Intermediate Algebra and Coordinate Geometry.** Items in this category are based on graphing in the standard coordinate plane or on other topics from intermediate algebra such as operations with integer exponents, radical expressions and rational expressions, the quadratic formula, linear inequalities in one variable, and systems of two linear equations in two variables.

d. **Plane Geometry.** Items in this category are based on the properties and relations of plane figures.

e. **Trigonometry.** Items in this category are based on right triangle trigonometry, graphs of the trigonometric functions, and basic trigonometric identities.

Table 2.3

Specifications for the ACT Reading Test

Description of the test. The Reading Test is a 40-item, 35-minute test that measures reading comprehension as a product of skill in referring and reasoning. That is, the test items require students to derive meaning from several texts by (1) referring to what is explicitly stated and (2) reasoning to determine implicit meanings and to draw conclusions, comparisons, and generalizations. The test comprises four prose passages (based on topics in prose fiction, the humanities, the social studies, and the natural sciences) that are representative of the level and kinds of writing commonly encountered in college freshman curricula. These passages are selected from published sources. Each passage is accompanied by a set of multiple-choice test items. These items do not test the rote recall of facts from outside the passage, isolated vocabulary items, or rules of formal logic. Rather, the test focuses upon the complex of complementary and mutually supportive skills that readers must bring to bear in studying written materials across a range of subject areas.

Three scores are reported for the ACT Reading Test: a total test score based on all 40 items, a subscore in Social Studies/Science reading skills (based on the 20 items in the social studies and natural sciences sections of the test), and a subscore in Arts/Literature reading skills (based on the 20 items in the prose fiction and humanities sections of the test).

Content of the test. The subject matter appropriate to each of the four types of reading selections and the approximate proportion of the test devoted to each are given below.

Reading Context	Proportion of Test	Number of Items
Social Studies	.25	10
Natural Sciences	.25	10
Prose Fiction	.25	10
Humanities	.25	10
Total	1.00	40

a. **Social Studies.** History, political science, economics, anthropology, psychology, sociology

b. **Natural Sciences.** Biology, chemistry, physics, physical sciences

c. **Prose Fiction.** Intact short stories or excerpts from short stories or novels

d. **Humanities.** Art, music, philosophy, theater, architecture, dance

BEST COPY AVAILABLE

Table 2.4

Specifications for the ACT Science Reasoning Test

Description of the test. The Science Reasoning Test is a 40-item, 35-minute test that measures the interpretation, analysis, evaluation, reasoning, and problem-solving skills required in the natural sciences. The test is made up of test units, each of which consists of some scientific information (the stimulus) and a set of multiple-choice test items. The scientific information is conveyed in one of three different formats: data representation, research summaries, or conflicting viewpoints.

The test items require students to recognize and understand the basic features of and concepts related to the provided information; to examine critically the relationships between the information provided and the conclusions drawn or hypotheses developed; and to generalize from given information to gain new information, draw conclusions, or make predictions.

A total test score only is reported for the ACT Science Reasoning Test.

Content of the test. The content of the Science Reasoning Test is drawn from Biology, Chemistry, Physics, and the Physical Sciences (e.g., Geology, Astronomy, and Meteorology). Advanced knowledge in these subjects is not required, but background knowledge at the level of a high school General Science course may be needed to answer some of the questions. Advanced mathematical skills are not required, but minimal arithmetic computations may be needed for some questions. Thus, the test emphasizes scientific reasoning skills rather than recall of scientific content, skill in mathematics, or reading ability. The approximate proportion of the test devoted to each of the three formats is given below.

Content Area*	Format	Proportion of Test	Number of Items
Biology Physical Sciences Chemistry Physics	Data Representation	.38	15
	Research Summaries	.45	18
	Conflicting Viewpoints	.17	7
Total		1.00	40

*Note: Content areas are distributed over the different formats.

a. **Data Representation.** This format presents students with graphic and tabular material similar to that found in science journals and texts. The items associated with this format measure skills such as graph reading, interpretation of scatter plots, and interpretation of information presented in tables. The graphic or tabular material may be taken from published materials; the items are composed expressly for the Science Reasoning Test.

b. **Research Summaries.** This format provides students with descriptions of one or more related experiments. The items focus upon the design of experiments and the interpretation of experimental results. The stimulus and items are written expressly for the Science Reasoning Test, and all relevant information is completely presented in the text of the stimulus or in the test questions.

c. **Conflicting Viewpoints.** This format presents students with expressions of several hypotheses or views that, being based on differing premises or on incomplete data, are inconsistent with one another. The items focus upon the understanding, analysis, and comparison of alternative viewpoints or hypotheses. Both the stimulus and the items are written expressly for the Science Reasoning Test.

BEST COPY AVAILABLE

Statistical Specifications

Statistical specifications for the tests indicate the level of difficulty (proportion correct) and minimum acceptable level of item discrimination (biserial correlation) of the test items to be used.

The distribution of item difficulties was selected so that the tests will effectively differentiate among students who vary widely in their level of achievement. Each of the four tests is constructed to have a mean item difficulty of about .58 for the ACT Assessment-tested population and a range of item difficulties from about .20 to .85.

With respect to discrimination indices, the standards for selecting items for the enhanced ACT Assessment tests are that (1) items in the difficulty range of .20-.29 should have a biserial correlation of .20 or higher with scores on a test measuring comparable content and (2) items in the difficulty range of .30-.85 should have a biserial correlation of .30 or higher with scores on a test measuring comparable content. Thus, for example, mathematics items would be evaluated with respect to students' performance on the ACT Assessment Mathematics Test.

Tests that satisfy these statistical specifications can be expected to yield a high degree of measurement precision for most students, while still providing good measurement for the high-achieving students who compete for scholarships.

Selection of Item Writers

Each year, ACT contracts with item writers to construct approximately 3,300 items for the enhanced ACT Assessment Program. The item writers are content specialists in the disciplines measured by the enhanced ACT Assessment tests. Most are actively engaged in teaching. They teach at a number of different levels, from high school to university, and at a variety of institutions, from small private institutions to large public ones. They represent the diversity of the population of the United States with respect to ethnic background, gender, and geographic location. ACT makes every effort to ensure that the item writers for the enhanced ACT Assessment Program represent a cross section of educators in the United States.

Before being asked to write items for the enhanced ACT Assessment tests, potential item writers are required to submit a sample set of materials for review. An item writer's guide that is specific to the content area is provided to each item writer. The guides include examples of items, as well as providing item writers with the test specifications and ACT's requirements with respect to content and style. Included are specifications for fair and representative portrayal of subgroups of individuals, avoidance of subject matter that may be unfamiliar to members of a subgroup, and nonsexist use of language.

Each unit submitted by a potential item writer is evaluated by ACT test development staff. A decision concerning whether to contract with the item writer is made on the basis of that evaluation.

Each item writer under contract is given an assignment to produce a small number of multiple-choice items. The size of the assignment ensures that a diversity of material will be produced and that the security of the testing program will be maintained, since any item writer will have knowledge of only a small proportion of the items produced. Item writers work closely with ACT test specialists, who assist them in producing items of high quality that meet the test specifications.

Item Construction

The item writers must create items that are educationally important as well as psychometrically sound. A large number of items must be constructed because, even with good writers, many items fail to meet the standards established for usable test items.

Each item writer submits a set of items, called a unit, in a content area. All items on the Mathematics Test are discrete (unrelated to a reading passage). All items on the English and Reading tests are related to reading passages, and all items on the Science Reasoning Test are related to some given scientific information.

Review of Items

After a unit is accepted, it is edited to meet ACT's specifications for content accuracy, word length, item classification, item format, and language. During the editing process, all test materials are reviewed for fair portrayal of subgroups and nonsexist use of language. The unit is reviewed numerous times by ACT staff to ensure that all of ACT's standards have been met.

A copy of each unit is then sent to consultants specializing in the appropriate content area for verification of content accuracy and a check of various quality-control considerations. These consultants have been previously trained by ACT staff. Any comments on the unit returned to ACT by the consultants are reviewed by ACT staff, and appropriate changes are made to the unit. Should there be disagreement with the consultants' suggestions concerning change in content, the unit is sent for review to different consultants. If any significant changes are made in the unit as a result of consultants' comments, those parts being changed are sent back to consultants to confirm that the changes were correctly made.

Item Tryouts

The items that are judged acceptable in the review process are assembled into tryout units for pretesting on samples from the national examinee population. These samples are carefully selected to be representative of the total examinee population. Each sample is administered a tryout unit from one of the four subject matter areas covered by the enhanced ACT Assessment tests. The time limits for the tryout units permit the majority of students to respond to all items.

Item Analysis of Tryout Units

Item analyses are performed on the tryout units. For a given unit the student sample is divided into groups by the individuals' scores on the enhanced ACT Assessment test in the same content area (taken at the same time as the tryout unit).

Proportions of students in each of these groups correctly answering each tryout item are tabulated, as well as the proportion in each group selecting each of the incorrect options. Biserial and point-biserial correlation coefficients between each item score (correct/incorrect) and the total score on the corresponding test of the regular (national) test form are also computed.

The item analyses serve to identify statistically effective test questions. Items that were either too difficult or too easy and those that failed to discriminate between students of high and low educational development, as measured by their corresponding enhanced ACT Assessment test scores, are eliminated or revised. The biserial and point-biserial correlation coefficients, as well as the differences between proportions of students answering the item correctly in each of the groups, are used as indices of the discriminating power of the pretest item.

Each item is reviewed following the item analysis. Items determined to be of poor quality are scrutinized by ACT staff in order to identify possible causes. Some items are revised and placed in new tryout units following the review. Additionally, the review process provides feedback that helps decrease the incidence of poor-quality items in the future.

Assembly of New Forms

Items that are judged acceptable in the review process are placed in an item pool. Preliminary forms of the enhanced ACT Assessment tests are constructed by selecting items from this pool that match both the content and statistical specifications for the tests. For each test in the battery, items for the new forms are selected to match the content distribution for the test as shown in Tables 2.1 through 2.4.

The preliminary versions of the test forms are subjected to several reviews to ensure that the items are accurate and that the overall test forms conform to good test construction practice. The first review is performed by the ACT staff. Items are checked for content accuracy and conformity to ACT style. The items are also reviewed to ensure that they are free of clues that could allow testwise students to answer the items correctly even though they lacked knowledge in the subject area.

The national forms are then sent to national consultants for review. For the English, Mathematics, and Reading tests, each form is reviewed by content consultants, measurement consultants, and minority consultants. For the Science Reasoning Test, each form is reviewed in its entirety by measurement consultants and minority consultants. In addition, content consultants review portions of the forms relevant to their areas of expertise (e.g., biology,

chemistry, physics, astronomy, or geology). These consultants are individuals different from those used for content reviews of tryout units.

In the final part of the review process, consultant panels are assembled for each of the four tests of the enhanced ACT Assessment. The panel for each test is made up of subject-matter experts from each relevant discipline as well as ACT staff. Since panel members discuss each item with reference to all pertinent disciplines as well as the test specifications, this procedure ensures content accuracy.

Comments from the consultants are then reviewed by ACT staff and any necessary changes are made. Whenever significant changes are made, the revised components are again reviewed by the appropriate consultants and by ACT staff. If additional content questions are raised, other consultants are asked to review the form. If no further corrections are needed, the test form is prepared for printing.

Review Following National Administration

Immediately after each national test administration, the data from each form of each test in the enhanced ACT Assessment are subjected to an item analysis. The results of these analyses are carefully scrutinized for such anomalies as substantial changes between pretest and national administrations in item difficulty and discrimination indices.

Occasionally, one or more examinees and/or an ACT test administrator raises a question about a particular test item after a test form is administered. During the review that follows each national administration, any item about which such a question has been raised is carefully examined by ACT staff and external content experts and a response detailing the results of the reviews is prepared and forwarded to the inquirer.

CHAPTER 3

NORMING, SCALING, AND EQUATING THE TESTS OF EDUCATIONAL DEVELOPMENT

Norms and Score Scale

In October 1988 a national study was conducted involving over 100,000 high school students in 399 schools. This study, called the Academic Skills Study, provided the data used to create the ACT score scales and provide nationally representative norms.

Sampling

The Academic Skills Study consisted of four samples, each intended to be nationally representative. Data from Sample 1 were used for the scaling of the enhanced ACT Assessment and to obtain nationally representative norms. Consequently, only Sample 1 is described here. A more detailed description of the Academic Skills Study and all four samples is given in Sawyer (in press).

In Sample 1 three test batteries (the enhanced ACT Assessment; a lengthened form of the P-ACT+, a tenth-grade program designed to complement the enhanced ACT Assessment; and a lengthened form of the original ACT Assessment) were administered to tenth-, eleventh-, and twelfth-grade students. To secure a random sample for each battery at each school, the stock of test booklets was arranged so that every third booklet was the same battery. The test administrators were instructed to distribute the booklets to the students without disturbing this order. This procedure is referred to as spiraling. Using spiraling, all batteries were administered at each school to randomly equivalent groups of examinees within each grade.

Sample Design and Data Collection. The target population consisted of students enrolled in tenth, eleventh, or twelfth grade in schools in the U.S. The target population included students in private as well as public schools.

The overall sampling design was two-stage sampling with stratification of first-stage units (schools). The explicit stratification variables used were: region of the country, school size, public versus private status, and ACT user versus non-ACT user. (A school was defined as an ACT user if it was located in the Mountain/Plains, Southwest, or Midwest regions or had at least one student who took the ACT Assessment in the 1987-88 academic year.) A systematic sample was selected from each stratum.

The minimum number of schools to be included in the sample was determined on the basis of the resulting precision in estimating percentiles. Given the sample design, it was estimated that at least 82 schools were needed to achieve the targeted precision, set as a 95% probability that $|\hat{p} - p| < .05$ for $p = .50$, where p is a percentile point of a score distribution for the national group of examinees, and $\hat{p}$ is the estimated value of p. This value of .05 is often referred to as the confidence interval half-width.

In anticipation that some schools would not participate in the study, many more schools were invited to participate than were required to achieve the targeted precision. During the recruitment the number of participating

schools in each stratum was carefully monitored, so as to maintain the representativeness of the sample with respect to the stratification variables. In addition, a backup sample was chosen so that additional schools could be chosen from strata for which there were too few schools agreeing to participate. Overall, 147 schools participated in the study.

Schools were asked to test all students in each grade. A few schools were allowed to administer the spiraled test batteries to randomly selected subsamples of their students. Makeup testing for students who were absent was strongly encouraged.

Data Editing. Data from one of the 147 schools that participated in the study were not used due to irregularities in the administration of the tests at that school. From the remaining 146 schools, examinees with problematic records were excluded (e.g., grade level not determinable, test form not determinable, zero raw score on one of the four tests, over two-thirds of the items omitted on any of the four tests). About 4% of returned answer sheets were excluded. The screening of examinees with problematic records resulted in all students from one school being excluded, thereby leaving 145 schools in the final sample. Final sample sizes for all examinees (national) and the subset of examinees who indicated they were college-bound are shown in Table 3.1. A college-bound student was defined as a student who answered "Yes" to the following question: "Are you planning to attend college after high school (either a 2-year community/junior college or a 4-year college/university)?"

Table 3.1

Examinee Sample Sizes

Grade	National	College-bound
10	5807	4980
11	5469	4659
12	5058	4306

Weighting. For the sampling and norming process, individual examinee records were multiplied by weights to represent the target population. The weight for an examinee is inversely proportional to the probability of the examinee being chosen for the sample, given the sampling plan. See Sawyer (in press) for details on the procedure used to determine the weights.

Response Rates. One type of nonresponse in this study was among schools: not every school invited to participate did so. Attempts were made to choose the replacement schools from the same strata as the schools they were replacing so that the obtained sample would be representative with respect to the stratification variables. Nevertheless, it is conceivable that schools' willingness to participate in this study could be related to their students' academic development, independently of these variables. If this were true, then the nonresponse among schools would introduce a bias in the results. However, given the results concerning sample representativeness presented subsequently, it is unlikely that the selection of schools had any significant biasing effect.

A second type of nonresponse was among students within a participating school. One source of student nonresponse was absenteeism (schools were encouraged to retest students who were absent). Within-school student participation rates are shown in Table 3.2 (results for schools who chose to test a subsample are not reported). The reported rates are the proportion of students actually tested relative to the number of enrolled students reported by the school. These data suggest that most schools tested nearly all of their students, as was requested, though at a few schools there were large differences between the number of students who were supposed to test and the number of students actually tested. If these low participation rates reflect some sort of selection related to test scores, then the sample from these schools may introduce bias in the results. It is believed that for the sample as a whole, student nonresponse has not had any important biasing effect.

Table 3.2

Within-School Student Participation Rates in Proportions

Grade	Minimum	Median	Maximum
10	.40	.93	1.00
11	.43	.92	1.00
12	.48	.90	1.00

Sample Representativeness. One way to determine the character and extent of sample bias is to compare the derographic characteristics of the sample of examinees with the U.S. statistics for various educational and demographic variables presented in Table 3.3. Precisely comparable U.S. data for the population of interest were not available. However, the data shown allow for a general examination of the representativeness of the sample with respect to the demographic variables.

As indicated in Table 3.3, the weighted sample appears to be reasonably representative of the national population. The largest apparent discrepancy is that the sample percentage of blacks is about 10% while the corresponding U.S. percentage is about 16%. However, the actual discrepancy is probably considerably less for two reasons. First, between 10% (grade 12) and 15% (grade 10) of the students in Sample 1 did not respond to the question concerning racial/ethnic background or chose "other" as their response. Second, the U.S. percentage of 16 in Table 3.3 is based on students in grades K-12, not just grades 10-12. To the extent that black students drop out at higher rates than other students, the U.S. percentage will be overstated.

The sample is also not completely geographically representative. At all three grade levels, there is a modest overrepresentation of students from the Southeast and an underrepresentation of students from the Northeast. This underrepresentativeness is a consequence of the relative levels of success in recruiting private schools from different parts of the country.

Table 3.3

Demographic and Educational Characteristics for Sample 1 of the Academic Skills Study

Category identifier used in study	Weighted sample proportion Grade 10	Grade 11	Grade 12	U.S. Proportion	U.S. category identifier
Sex					
Female	.49	.50	.50	.50	Female
Male	.50	.50	.50	.50	Male
Racial/ethnic					
Afro-American/Black	.10	.10	.10	.16	Black
American Indian, Alaskan Native	.02	.02	.01	.01	Indian
Caucasian American/White	.68	.71	.74	.71	White
Mexican American/Chicano	.03	.02	.02	.09	Spanish Origin
Puerto Rican, Cuban, Other Hispanic	.01	.01	.01	--	--
Asian American/Pacific Islander	.02	.02	.02	.03	Asian
Other, Prefer Not to Respond, Blank	.15	.12	.10	--	--
School affiliation					
Private	.10	.11	.11	.09	Private
Public	.90	.89	.89	.91	Public
Region					
West Coast	.15	.15	.15	.16	West Coast
Mountains/Plains	.07	.07	.07	.06	Mtns/Plains
Southwest	.12	.12	.11	.12	Southwest
Midwest	.22	.23	.24	.25	Midwest
Southeast	.26	.24	.24	.19	Southeast
Northeast	.18	.18	.18	.22	Northeast

Obtained Precision. Because data from more schools were available (145) than were targeted (82), the obtained precision was considerably better than the targeted level. The targeted 95% confidence interval half-width for $p = .50$ was the value .05. The median obtained value was .03 for grades 10, 11, and 12.

Scaling

Scale scores are reported for the enhanced ACT Assessment English, Mathematics, Reading, and Science Reasoning tests, and for the Usage/Mechanics, Rhetorical Skills, Pre-Algebra/Elementary Algebra, Intermediate Algebra/Coordinate Geometry, Plane Geometry/Trigonometry, Social Studies/Science Reading, and Arts/Literature Reading subscores. A Composite score, calculated by rounding the unweighted average of the four test scores, is also reported. The rounding is done such that if the fractional part of the average is 0 or .25, the integer part of the average is used as the Composite score. If the fractional part of the average is .50 or .75, the integer part of the average plus 1 is used as the Composite score. For example, if the average of the four scores is 17.5, a Composite score of 18

will be reported. There is no arithmetic relationship between subscores and test scores, because subscores and test scores were all scaled separately.

The Score Scale. The range of the test scores and Composite scores on the enhanced ACT Assessment is 1 to 36. The range of the subscores is 1 to 18. Although the score scale for the enhanced ACT Assessment and the score scale for the original ACT Assessment that was begun in 1959 have the same score points, scale scores on these two tests are *not* directly comparable due to changes in the internal structure of the tests and the methodology used for scaling. The means of the enhanced ACT Assessment score scales are intended to be 18 for each of the four tests and the Composite and 9 for the seven subscores among students at the beginning of twelfth grade, nationwide, who reported that they plan to attend a two- or four-year college.

For the enhanced ACT Assessment, the standard error of measurement was targeted at approximately 2 scale score points for each of the test scores and subscores and 1 scale score point for the Composite. In addition, the scales for the enhanced ACT Assessment were constructed using a method described by Kolen (1988) to produce score scales with approximately equal standard errors of measurement along the entire range of scores. Without nearly equal standard errors of measurement, standard errors of measurement at different score levels would need to be presented and considered in score interpretation (see Standards for Educational and Psychological Testing, APA, 1985, p. 22). Based on the properties just described, if the distribution of measurement error is approximated by a normal distribution, an approximate 68% confidence interval can be constructed for any examinee by adding 2 points to and subtracting 2 points from his or her scale score for any of the enhanced ACT Assessment tests or subscores. An analogous interval for the Composite can be constructed by adding 1 point to and subtracting 1 point from the Composite.

In thinking about standard errors and their use, note that the reported scale score (i.e., the obtained score) for an examinee is only an estimate of that examinee's true score, where the true score can be interpreted as the average obtained score over repeated testings under identical conditions. If 1 standard error of measurement were added to and subtracted from each of these obtained scores, about 68% of the resulting intervals would contain the examinee's true score. Technically, this is how a 68% interval for an examinee's scale score should be interpreted. These statements make normal distribution assumptions.

Another way to view 68% intervals is in terms of groups of examinees. Specifically, if 1 standard error of measurement were added to and subtracted from the obtained score of each examinee in a group of examinees, the resulting intervals would contain the true score for approximately 68% of the examinees. To put it another way, about 68% of the examinees would be mismeasured by less than 1 standard error of measurement. Again, such statements make normal distribution assumptions. Also, these statements assume a constant standard error of measurement, which is a characteristic (approximately) of the enhanced ACT Assessment score scales. Consequently, it is relatively straightforward to interpret scale scores in relation to measurement error.

Properties of the score scale for the enhanced ACT Assessment are summarized in Table 3.4.

Table 3.4

Properties of the Score Scale for the Enhanced ACT Assessment

- Score scale is new. Scores on the enhanced ACT Assessment and the original ACT Assessment are not directly comparable.
- Range of scores is 1-36 on all tests and the Composite, and 1-18 for subscores.
- Test means are 18 and subscore means are 9 among twelfth-grade U.S. students who report they plan to attend college.
- Standard error of measurement is approximately 2 points for each test score and subscore and 1 point for the Composite. This property is used to discourage overinterpretation of small score differences. Approximately 68% of the examinees will have observed scores within 2 points of their true scores on the tests and within 1 point of their true score on the Composite.
- The standard error of measurement is approximately equal along the score scale.
- The occurrence of gaps (unused scale score points) and multiple raw scores converting to the same scale score were minimized in constructing the raw-to-scale score transformation.

Scaling Process. The scaling process for the enhanced ACT Assessment consisted of three steps. First, weighted raw score distributions for both national and college-bound groups of examinees from the Academic Skills Study were computed, with the weighting based on the sample design. Second, the weighted raw score distributions were smoothed according to a four parameter beta compound binomial model (Kolen & Hanson, in press; Lord, 1965). Finally, the smoothed college-bound twelfth-grade raw score distributions were used to produce the score scales.

The purpose of smoothing the raw score distributions was to produce distributions that are easier to work with and that are better estimates of population distributions. Cope and Kolen (1987) and Hanson (1988) have shown that smoothing techniques have the potential to improve the estimation of population distributions. Overall, the smoothing process resulted in distributions that appeared smooth without departing too much from the unsmoothed distributions. In addition, the first three central moments (mean, standard deviation, skewness) of the smoothed distributions were identical to those of the original distributions. The fourth central moments of the smoothed distributions (kurtosis) were either identical or very close to those of the original distributions.

The first step in constructing the score scales was to produce initial scale scores with a specified mean for twelfth-grade college-bound examinees from the Academic Skills Study and a specified standard error of measurement that is approximately equal for all examinees. The means and standard errors of measurement specified for each test score and subscore were those given in Table 3.4. The process used was based on Kolen (1988) and described in detail by Kolen and Hanson (in press). These initial scale scores were rounded to integers ranging from 1 to 36 for the tests and 1 to 18 for the subscores. Some adjustment of the rounded scale scores was performed to attempt to meet the specified mean and standard error of measurement, as well as to avoid gaps in the score scale (scale scores that were not used) or to avoid having too many raw scores converting to a single scale score. This process resulted in the final raw-to-scale score conversions.

Scale score summary statistics, standard errors of measurement (SEM), and reliabilities for twelfth-grade students are given in Table 3.5 for the test scores and the Composite scores and in Table 3.6 for the subscores. The mean scale scores for college-bound students were, in all cases, within .01 of the target values given in Table 3.4. The Composite mean is somewhat higher than 18 due to the rounding rule used to form the Composite from the test scores. The variation in the standard errors of measurement among the various test scores and subscores was a result of differences in raw score means, reliabilities, and test lengths.

Table 3.5

Scale Score Summary Statistics for Enhanced ACT Assessment Tests

Statistic	English	Mathematics	Reading	Science Reasoning	Composite
National					
Mean	17.18	17.44	17.17	17.48	17.45
Standard Deviation	5.29	4.44	6.40	4.32	4.54
Skewness	0.59	0.98	0.60	0.78	0.78
Kurtosis	2.84	3.74	2.79	3.69	3.21
SEM	1.50	1.50	2.21	2.04	0.92
Reliability	0.92	0.89	0.88	0.78	0.96
College-bound					
Mean	18.01	18.01	17.99	17.99	18.13
Standard Deviation	5.25	4.54	6.45	4.40	4.56
Skewness	0.47	0.85	0.49	0.72	0.66
Kurtosis	2.71	3.41	2.65	3.51	3.03
SEM	1.50	1.49	2.21	2.03	0.92
Reliability	0.92	0.89	0.88	0.79	0.96

Table 3.6

Scale Score Summary Statistics for Subscores of Enhanced ACT Assessment

	Usage/ Mechanics	Rhetorical Skills	Pre-Algebra/ Elementary Algebra	Intermediate Algebra/ Coordinate Geometry	Plane Geometry/ Trigonometry	Reading Social Studies/ Sciences	Reading Arts/ Literature
National							
Mean	8.50	8.57	8.57	8.76	8.73	8.68	8.40
SD	3.43	2.99	3.30	2.71	2.85	3.19	4.52
Skewness	0.50	0.51	0.46	0.06	-0.05	0.59	0.37
Kurtosis	2.73	2.90	2.63	3.66	3.27	2.87	2.09
SEM	1.27	1.23	1.42	1.87	1.80	1.49	1.89
Reliability	0.86	0.83	0.81	0.52	0.60	0.78	0.83
College-Bound							
Mean	9.01	9.01	9.01	9.00	9.01	8.99	9.01
SD	3.40	2.98	3.31	2.76	2.87	3.27	4.53
Skewness	0.40	0.40	0.34	0.05	-0.10	0.49	0.22
Kurtosis	2.63	2.78	2.53	3.56	3.23	2.69	2.00
SEM	1.26	1.23	1.42	1.84	1.75	1.49	1.91
Reliability	0.86	0.83	0.82	0.55	0.63	0.79	0.82

Norming

National Study Norms. The norms for the enhanced ACT Assessment were obtained by the sampling and weighting process described previously, and are intended to represent the national population of twelfth-grade high school students and the national subpopulation of twelfth-grade high school students who reported that they were college-bound.

Data from Sample 1 of the Academic Skills Study were used to develop cumulative percents for each enhanced ACT Assessment test and the Composite, as well as the subscores. The cumulative percent corresponding to a given scale score is defined as the percent of examinees with scores equal to or less than that score. Tables 3.7 and 3.8 contain cumulative percents for the four enhanced ACT Assessment test scores and the Composite score for college-bound twelfth-grade students and national twelfth-grade students, respectively. Tables 3.9 and 3.10 contain cumulative percents for the seven enhanced ACT Assessment subscores for college-bound twelfth-grade students and national twelfth-grade students, respectively.

An examinee's standing on different tests should be compared by using the cumulative percents shown in the norms tables and not by using scale scores. The reason for preferring cumulative percents for such comparisons is that the scales were not constructed to ensure that, for example, a scale score of 6 on the English Test is comparable to a 6 on the Mathematics or Reading or Science Reasoning Test. In contrast, examinee cumulative percents on different tests indicate standings relative to the same comparison group (either college-bound or national).

Table 3.7

Estimated Cumulative Percents on Enhanced ACT Assessment Tests for Fall College-Bound Twelfth-Grade Students

Scale Score	English	Mathematics	Reading	Science Reasoning	Composite
01	01	01	01	01	01
02	01	01	01	01	01
03	01	01	01	01	01
04	01	01	01	01	01
05	01	01	01	01	01
06	01	01	01	01	01
07	01	01	02	01	01
08	01	01	03	01	01
09	02	01	08	01	01
10	05	01	12	02	01
11	09	02	16	04	03
12	16	07	21	07	08
13	23	14	30	11	16
14	28	24	34	23	24
15	36	33	39	30	32
16	43	46	47	44	41
17	50	53	52	51	50
18	57	62	55	62	59
19	64	69	63	66	66
20	70	74	66	75	73
21	75	80	73	81	78
22	79	83	76	84	83
23	84	87	79	89	87
24	88	90	82	91	89
25	90	93	87	94	92
26	93	94	89	95	95
27	95	96	91	97	96
28	96	97	93	98	98
29	98	98	94	98	98
30	99	99	96	99	99
31	99	99	97	99	99
32	99	99	97	99	99
33	99	99	98	99	99
34	99	99	99	99	99
35	99	99	99	99	99
36	99	99	99	99	99

Table 3.8

Estimated Cumulative Percents on Enhanced ACT Assessment Tests for Fall National Twelfth-Grade Students

Scale Score	English	Mathematics	Reading	Science Reasoning	Composite
01	01	01	01	01	01
02	01	01	01	01	01
03	01	01	01	01	01
04	01	01	01	01	01
05	01	01	01	01	01
06	01	01	01	01	01
07	01	01	02	01	01
08	01	01	04	01	01
09	03	01	11	01	01
10	08	01	15	02	01
11	13	03	20	05	04
12	21	09	25	09	11
13	30	18	35	14	22
14	35	29	40	28	31
15	43	39	45	35	40
16	50	52	53	50	48
17	57	59	57	56	57
18	64	67	61	66	65
19	69	74	68	71	71
20	75	78	71	78	77
21	79	83	77	84	81
22	82	86	80	86	85
23	86	89	82	91	89
24	90	91	84	92	91
25	92	94	89	95	93
26	94	95	91	96	96
27	96	97	92	98	97
28	97	98	94	98	98
29	98	98	95	99	99
30	99	99	97	99	99
31	99	99	98	99	99
32	99	99	98	99	99
33	99	99	99	99	99
34	99	99	99	99	99
35	99	99	99	99	99
36	99	99	99	99	99

Table 3.9

Estimated Cumulative Percents for Enhanced ACT Assessment Subscores for Fall College-Bound Twelfth-Grade Students

Scale Score	Usage/ Mechanics	Rhetorical Skills	Pre-Algebra/ Elementary Algebra	Intermediate Algebra/ Coordinate Geometry	Plane Geometry/ Trigonometry	Reading Social Studies/ Sciences	Reading Arts/ Literature
01	01	01	01	01	01	01	01
02	01	01	01	01	01	01	06
03	02	02	02	04	05	02	10
04	07	04	09	04	05	06	22
05	14	09	14	13	14	13	28
06	27	23	27	13	14	22	35
07	36	33	34	26	27	41	42
08	49	49	47	42	40	49	48
09	57	59	59	57	62	64	55
10	68	72	70	76	70	70	62
11	78	80	75	86	82	75	68
12	84	86	84	89	90	84	75
13	88	91	88	94	93	88	81
14	92	95	94	96	97	94	87
15	95	98	97	99	99	96	87
16	98	99	98	99	99	98	93
17	99	99	99	99	99	99	98
18	99	99	99	99	99	99	99

Table 3.10

Estimated Cumulative Percents for Enhanced ACT Assessment Subscores for Fall National Twelfth-Grade Students

Scale Score	Usage/ Mechanics	Rhetorical Skills	Pre-Algebra/ Elementary Algebra	Intermediate Algebra/ Coordinate Geometry	Plane Geometry/ Trigonometry	Reading Social Studies/ Sciences	Reading Arts/ Literature
01	01	01	01	01	01	01	01
02	02	01	01	01	01	01	08
03	04	03	02	05	06	02	13
04	10	05	11	05	06	06	27
05	19	13	18	14	16	14	34
06	34	29	33	14	16	24	41
07	43	40	40	29	30	45	48
08	56	55	53	46	44	54	54
09	63	65	65	60	66	68	61
10	73	76	75	79	74	73	67
11	81	83	79	88	85	78	73
12	86	88	87	91	92	87	79
13	90	93	90	95	94	90	84
14	94	96	95	97	98	95	90
15	96	98	97	99	99	97	90
16	98	99	99	99	99	98	94
17	99	99	99	99	99	99	98
18	99	99	99	99	99	99	99

Even comparisons of cumulative percents do not permit comparisons of standing in different skill areas in any absolute sense. The question of whether a particular examinee is stronger in Science Reasoning than in Mathematics can be answered only in relation to reference groups of other examinees. Whether the answer is "yes" or "no" can depend on the group.

Additional Norms Reported for 1989-90. In addition to reporting national college-bound norms based on the Academic Skills Study, estimated norms based on ACT-tested high school graduates of 1986, 1987, and 1988, for the English, Mathematics, and Composite scores are also reported. These ACT Assessment high school graduate norms, shown in Table 3.11, were determined by applying concordance tables (see Appendix B) to link original ACT Assessment scores to enhanced ACT Assessment scores. The concordance table, therefore, makes it possible for reports prepared in 1989-90 to report enhanced ACT Assessment norms based on previous years' ACT-tested high school graduates.

Equating

Several new forms of each of the enhanced ACT Assessment tests are developed each year. Even though each form is constructed to adhere to the same content and statistical specifications, the forms may be slightly different in difficulty. To control for these differences, subsequent forms are equated, and the scores reported to examinees are scale scores that have the same meaning regardless of the particular form administered to examinees. Thus, scale scores are comparable across test forms and test dates.

A carefully selected sample of examinees from each year's October national test date is used as an equating sample. The examinees in this sample are administered a spiraled set of forms--the new forms ("n-1" of them) and one anchor form that has already been equated to previous forms. (Initially, of course, the anchor form is the form used to establish the score scale.) This spiraling technique, in which every nth examinee receives the same form of the test, results in randomly equivalent groups taking the forms. The use of randomly equivalent groups is an important feature of the equating procedure and provides a basis for a great deal of confidence in the continuity of scales. More than 2,000 examinees take each form.

Scores on the alternate forms are equated to the score scale using equipercentile equating methodology. In equipercentile equating, a score on Form X of a test and a score on Form Y are considered to be equivalent if they have the same percentile rank in a given group of examinees. The equipercentile equating results are subsequently smoothed using an analytic method described by Kolen (1984) to establish a smooth curve, and the equivalents are rounded to integers. The conversion tables that result from this process are used to transform raw scores on the new forms to scale scores.

The equipercentile equating technique is applied to the raw scores of each of the four tests for each form separately. The Composite score is not directly equated across forms. Instead, it is a rounded arithmetic average of

Table 3.11

National Percentile Distributions of ACT Assessment Scale Scores, 1986-87-88 ACT-Tested High School Graduating Classes

Scale Score	English			Mathematics			Social Studies			Natural Sciences			Composite			Scale Score
	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T	
36	99	99	99	99	99	99	0	0	0	0	0	0	99	99	99	36
35	99	99	99	99	99	99	0	0	0	0	0	0	99	99	99	35
34	99	99	99	98	99	99	0	0	0	0	0	0	99	99	99	34
33	99	99	99	97	99	98	0	0	0	0	0	0	99	99	99	33
32	99	98	98	96	99	98	0	0	0	0	0	0	99	99	99	32
31	97	96	97	94	98	96	0	0	0	0	0	0	98	99	99	31
30	95	94	95	92	97	95	0	0	0	0	0	0	97	99	98	30
29	93	91	92	89	95	92	0	0	0	0	0	0	94	97	96	29
28	91	88	90	89	95	92	0	0	0	0	0	0	91	95	93	28
27	87	83	85	85	93	89	0	0	0	0	0	0	87	93	90	27
26	80	76	78	79	88	84	0	0	0	0	0	0	82	89	86	26
25	80	76	78	74	85	80	0	0	0	0	0	0	77	85	81	25
24	73	67	70	70	81	76	0	0	0	0	0	0	71	81	76	24
23	65	58	61	70	81	76	0	0	0	0	0	0	65	76	71	23
22	65	58	61	64	76	70	0	0	0	0	0	0	59	70	65	22
21	56	49	52	54	66	60	0	0	0	0	0	0	53	64	59	21
20	50	42	46	46	59	53	0	0	0	0	0	0	42	52	48	20
19	43	35	39	42	54	49	0	0	0	0	0	0	37	46	42	19
18	36	29	32	35	46	41	0	0	0	0	0	0	27	34	31	18
17	30	23	26	29	39	35	0	0	0	0	0	0	22	29	26	17
16	25	19	22	25	35	30	0	0	0	0	0	0	14	19	17	16
15	21	16	18	16	24	20	0	0	0	0	0	0	11	14	13	15
14	13	9	11	8	13	11	0	0	0	0	0	0	8	10	9	14
13	11	7	9	5	7	6	0	0	0	0	0	0	4	4	4	13
12	6	4	5	2	4	3	0	0	0	0	0	0	2	2	2	12
11	3	2	3	1	3	2	0	0	0	0	0	0	1	1	1	11
10	2	1	2	1	1	1	0	0	0	0	0	0	1	1	1	10
9	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	9
8	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	8
7	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	7
6	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	6
5	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	5
4	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	4
3	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	3
2	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	2
1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1
Mean	20.8	21.8	21.3	21.4	19.7	20.5	0.0	0.0	0.0	0.0	0.0	0.0	21.6	20.4	20.9	
S.D.	5.6	5.5	5.6	5.7	5.1	5.4	0.0	0.0	0.0	0.0	0.0	0.0	4.8	4.5	4.7	

Sample size Male (M): 1,078,625 Female (F): 1,272,860 Total (T): 2,351,485

These three-year norms are the source of norms provided on ACT Assessment score reports during the 1989-90 testing year. The norms are based on high school graduates who elected to take the ACT Assessment.

the scale scores for the four equated tests. The subscores are also separately equated using the equipercentile method. Note, in particular, that the equating process does *not* lead to a given test score being a prespecified arithmetic combination of subscores.

CHAPTER 4

RELIABILITY AND VALIDITY OF THE TESTS OF EDUCATIONAL DEVELOPMENT

Reliability, Measurement Error, and Effective Weights

Some degree of inconsistency or error is potentially contained in the measurement of any cognitive characteristic. An examinee administered one form of a test on one occasion and a second, parallel form on another occasion likely would earn somewhat different scores on the two administrations. These differences might be due to the examinee or the testing situation, such as differential motivation or differential levels of distractions on the two testings. Alternatively, these differences might result from attempting to infer the examinee's level of skill from a relatively small sample of items.

Reliability coefficients are estimates of the consistency of test scores. They typically range from zero to one, with values near one indicating greater consistency and those near zero indicating little or no consistency.

The standard error of measurement (SEM) is closely related to test reliability. The standard error of measurement summarizes the amount of error or inconsistency in scores on a test. As noted previously, the score scales for the enhanced ACT Assessment were developed to have approximately constant standard errors of measurement for all scale scores (i.e., the conditional standard error of measurement as a function of true scale score is constant). This statement implies, for example, that the standard error of measurement for any particular enhanced ACT Assessment test score or subscore is approximately the same for low-scoring examinees as it is for high-scoring examinees. As discussed more fully in a previous section, if the distribution of measurement error is approximated by a normal distribution, about two-thirds of the examinees can be expected to be mismeasured by less than 1 standard error of measurement.

Figure 4.1 presents the standard errors of measurement for the four tests as a function of true scale score. The minimum true scale score plotted is around 10 for each test because there was only a very small proportion of examinees with a ***true*** scale score this low for any of the tests. See Kolen and Hanson (in press) for details on the method used to produce these plots.

For most of the true scale score range, the scale score standard error of measurement is reasonably constant for each of the four tests. However, the standard error of measurement is smaller at high scores. This smaller standard error of measurement occurs at very high true scale scores (around 33) for the English Test, and at somewhat lower scores for the other tests. The smaller standard error of measurement resulted primarily from the need to truncate scores near the top to avoid gaps near the upper end of the raw-to-scale score conversion tables. An additional reason is that the method used to produce the score scales cannot guarantee a completely constant standard error of measurement for all true scale scores.

Enhanced ACT Assessment reliability coefficients and standard errors of measurement for the four tests and Composite are shown in Table 4.1 for the national and college-bound groups of examinees. Reliability coefficients and

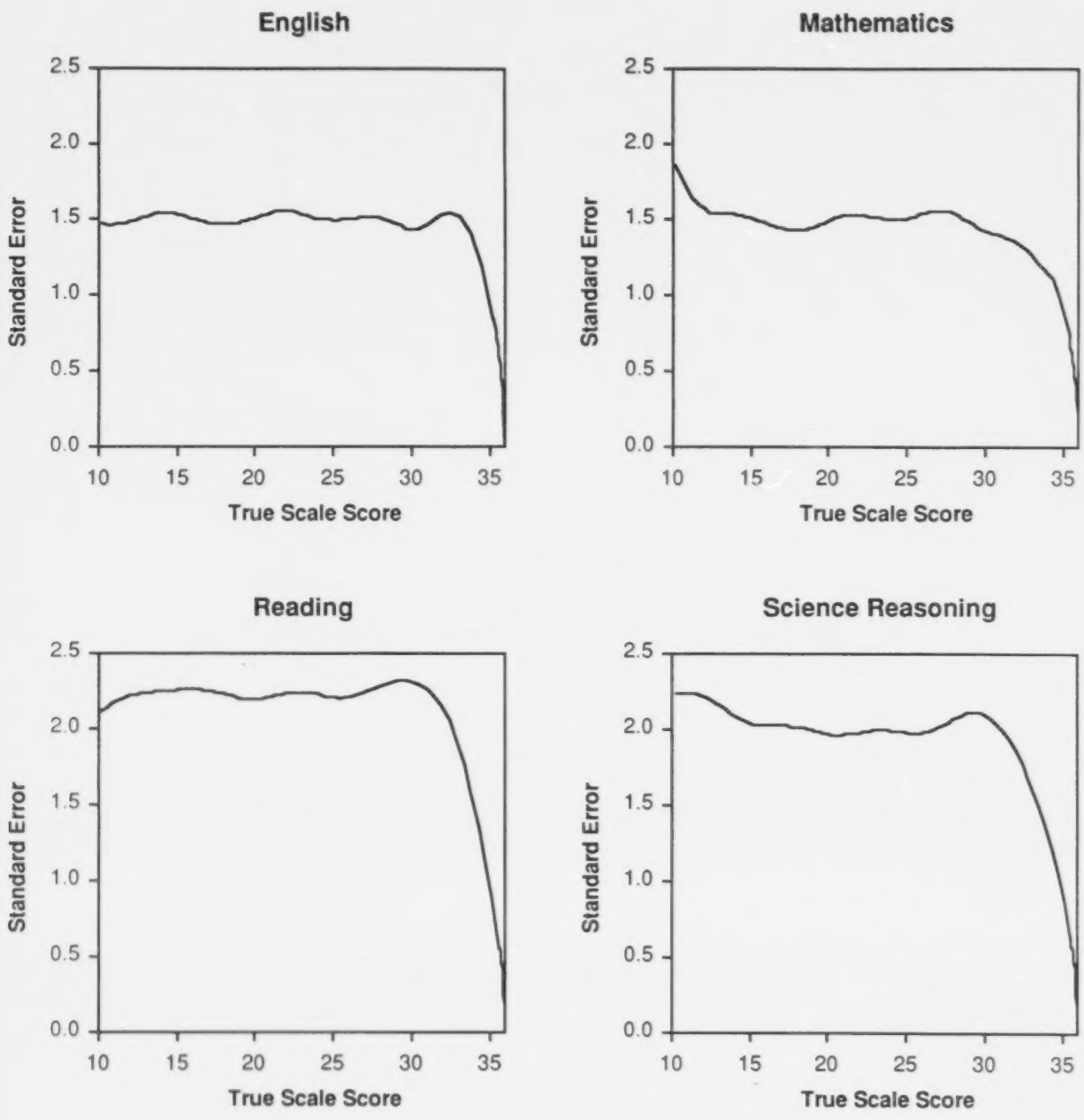

Figure 4.1. Conditional Standard Error of Measurement for Enhanced ACT Assessment Tests.

Table 4.1

Estimated Reliabilities and Standard Errors of Measurement for Enhanced ACT Assessment Tests

Statistic	English	Mathematics	Reading	Science Reasoning	Composite
National					
KR20 (raw scores)	0.93	0.90	0.89	0.81	---
Scale Scores					
Reliability	0.92	0.89	0.88	0.78	0.96
SEM	1.50	1.50	2.21	2.04	0.92
College-Bound					
KR20 (raw scores)	0.93	0.90	0.89	0.82	---
Scale Scores					
Reliability	0.92	0.89	0.88	0.79	0.96
SEM	1.50	1.49	2.21	2.03	0.92

standard errors of measurement for the subscores are shown in Table 4.2. In all cases the data used to estimate the reliabilities and standard errors of measurement are grade-twelve examinees from Sample 1 of the Academic Skills Study. Kuder-Richardson 20 (KR20) internal consistency reliability coefficients of raw scores are listed first. Scale score reliability coefficients and standard errors of measurement are reported next. Scale score standard errors of measurement were estimated using a four parameter beta compound binomial model as described by Kolen and Hanson (in press). The estimated scale score reliability, described by Kolen and Hanson (in press), was calculated for test i (REL_i) as

$$REL_i = 1 - \frac{SEM_i^2}{s_i^2} ,$$

where SEM_i is the estimated scale score standard error of measurement and s_i^2 is the observed scale score variance for test i.

The estimated standard error of measurement for the Composite (SEM_c) was calculated as

$$SEM_c = \tfrac{1}{4} \left[\sum_i SEM_i^2\right]^{\frac{1}{2}} ,$$

where the summation is over the four tests. The estimated reliability of the Composite (REL_c) was calculated as

$$REL_c = 1 - \frac{SEM_c^2}{s_c^2} ,$$

where s_c^2 is the estimated variance of scale scores on the Composite.

Table 4.2

Estimated Reliabilities and Standard Errors of Measurement for Enhanced ACT Assessment Subscores

	Usage/ Mechanics	Rhetorical Skills	Pre-Algebra/ Elementary Algebra	Intermediate Algebra/ Coordinate Geometry	Plane Geometry/ Trigonometry	Reading Social Studies/ Sciences	Reading Arts/ Literature
National							
Raw scores							
KR20	0.89	0.85	0.84	0.70	0.73	0.81	0.83
Scale Scores							
Reliability	0.86	0.83	0.81	0.52	0.60	0.78	0.83
SEM	1.27	1.23	1.42	1.87	1.80	1.49	1.89
College-Bound							
Raw scores							
KR20	0.89	0.85	0.84	0.72	0.74	0.82	0.83
Scale Scores							
Reliability	0.86	0.83	0.82	0.55	0.63	0.79	0.82
SEM	1.26	1.23	1.42	1.84	1.75	1.49	1.91

Scale scores from the four tests are summed together and divided by 4 in the process of calculating the Composite score. This process suggests that, in a sense, each test is contributing equally to the Composite. The weights used ($\frac{1}{4}$, in this case) are often referred to as ***nominal weights***.

Other definitions of the contribution of a test to a composite may be more useful. Wang and Stanley (1970) described ***effective weights*** as an index of the contribution of a test to a composite (see also Petersen, Kolen, & Hoover, 1988, and Jarjoura & Brennan, 1982). Specifically, the effective weight is defined as the covariance between a test score and the score on a composite. These covariances can be summed over tests and then each covariance divided by their sum (i.e., the composite variance) to arrive at ***proportional effective weights***. Proportional effective weights are referred to as effective weights in the remainder of this discussion.

The covariances and effective weights based on twelfth-grade examinees from the Academic Skills Study are shown in Table 4.3. The values in the diagonals that are not in parentheses are the observed scale score variances. With nominal weights of $\frac{1}{4}$ for each test, the effective weight for a test can be calculated by summing the values in the appropriate row in Table 4.3 that are not in parentheses and dividing the resulting value by the sum of all covariances using the formula

$$(\text{effective weight})_i = \frac{\sum_j \text{cov}_{ij}}{\sum_i \sum_j \text{cov}_{ij}},$$

where cov_{ij} is the observed covariance of test scores corresponding to row i and column j in Table 4.3. Effective weights for true scores, shown in

parentheses, are calculated similarly, with the true score variance $[s_i^2 \cdot REL_i]$ used in place of the observed score variance.

The effective weights for the national and college-bound groups are virtually identical. The effective weight for Reading is the largest of the effective weights. It is relatively high because Reading had the largest scale score variance, and because its covariances with the other measures tended to be highest. English has the next highest effective weight. These effective weights imply that the Reading and English tests are more heavily weighted (relative to Composite variance) in forming the Composite than are the Mathematics and Science Reasoning tests. Note that these effective weights are for the nationally representative sample obtained from the Academic Skills Study, and the weights might differ considerably from these for other examinee groups.

Table 4.3

Scale Score Covariances and Effective Weights

	English	Mathematics	Reading	Science Reasoning
No. of Items	75	60	40	40
Proportion of Total Items	.35	.28	.19	.19
National Scale Score				
English	28.10(25.85)	16.98	26.00	15.84
Mathematics	16.98	19.82(17.64)	18.13	12.69
Reading	26.00	18.13	41.19(36.25)	20.55
Science Reasoning	15.84	12.69	20.55	18.83(14.69)
Effective Weight	.26(.27)	.21(.21)	.32(.32)	.21(.20)
Reliability	.92	.89	.88	.78
College-Bound Scale Score				
English	27.72(25.50)	16.99	25.63	15.80
Mathematics	16.99	20.75(18.46)	18.25	13.13
Reading	25.63	18.25	41.82(36.80)	21.00
Science Reasoning	15.80	13.13	21.00	19.46(15.37)
Effective Weight	.26(.26)	.21(.21)	.32(.32)	.21(.21)
Reliability	.92	.89	.88	.79

Note: Numbers in parentheses relate to true scores (see Jarjoura & Brennan, 1982).

Validity

The primary validity evidence currently available for the enhanced ACT Assessment relates to content validity, and much of this evidence is given in the descriptions of the tests that were presented earlier in this manual. However, because the basic philosophy of the ACT Assessment has not changed, the validity data available for the original ACT Assessment may be of interest, and may be found in the ACT Assessment Program Technical Manual

(ACT, 1988). Because the enhanced ACT Assessment is a new program that, at this writing, has not been administered operationally, little statistical data on criterion-related validity are currently available. Validity studies are currently being planned and implemented, including those to study relationships between enhanced ACT Assessment scores and college grade point average.

The limited statistical validity evidence currently available is in the form of relationships among test scores (correlations and disattenuated correlations) and differential test performance of groups with different educational backgrounds using data from the Academic Skills Study.

Content Validity

The fundamental idea underlying development of the four tests is that the best way to predict success in college is to measure as directly as possible the academic skills developed in high school that are important for success in college. This means the tasks presented in the tests must be representative of scholastic tasks. They must be intricate in structure; they must be comprehensive in scope; and they must be significant in their own right, rather than narrow or artificial tasks that can be defended for inclusion only on the grounds of their statistical correlation with a criterion.

The enhanced ACT Assessment tests contain a large proportion of complex problem-solving exercises and proportionately few measures of narrow skills. They are oriented toward major areas of college and high school instructional programs rather than toward a factorial definition of various aspects of intelligence. Thus, the scores have a direct and obvious relationship to the student's educational progress, and a meaning that can be readily grasped by the instructional staff, the parents, and the student.

Since the tests are designed to measure as directly as possible the degree to which each student has developed in general skills and abilities needed for success in college work, content-related validity is of particular significance.

As described in Chapter 2, the enhanced ACT Assessment test-development procedures include an extensive review process, with each item being critically examined at least twelve times. Detailed test specifications have been developed to ensure that test content is representative of current high school and university curricula. All forms are reviewed to ensure that they match these specifications. A major component of the review process is a review of the content of the tests assessed. Hence, there is an ongoing assessment of the content validity of the tests during the test development process.

Test Correlations

Table 4.4 displays the intercorrelations between the scale scores on pairs of tests and the Composite for twelfth-grade students for both the college-bound and national groups. Values above the diagonal are observed correlations, and values below the diagonal are disattenuated correlations. Disattenuated correlations (observed correlations corrected for unreliability) can be interpreted as an estimate of the correlation between tests that are

perfectly reliable (i.e., true score correlations). The disattenuated correlation between scores on tests i and j was computed as

$$\text{disattenuated}(r_{ij}) = \frac{r_{ij}}{\sqrt{REL_i REL_j}},$$

where r_{ij} is the observed correlation between scores on tests i and j.

As expected, the values in Table 4.4 illustrate that the disattenuated correlation between two tests is greater than the observed correlation. The scale scores on the four tests have correlations for college-bound students in the interval .60-.75, indicating that examinees who score well on one test also tend to score well on another. The disattenuated correlations among tests shown in Table 4.4 are sufficiently below 1.0 to suggest that the tests are measuring skills that are at least somewhat distinct, statistically. The disattenuated correlation between Science Reasoning and Reading is the highest of the disattenuated correlations between tests shown in the table and the disattenuated correlation between Mathematics and Reading is the lowest.

Table 4.4

Observed Correlations (Above Diagonal) and Disattenuated Correlations (Below Diagonal) Between Enhanced ACT Assessment Test Scores

	English	Mathematics	Reading	Science Reasoning	Composite
National					
English	1.00	.72	.76	.69	.90
Mathematics	.80	1.00	.63	.66	.84
Reading	.84	.71	1.00	.74	.91
Science Reasoning	.81	.79	.89	1.00	.86
College-Bound					
English	1.00	.71	.75	.68	.90
Mathematics	.78	1.00	.62	.65	.83
Reading	.83	.70	1.00	.74	.91
Science Reasoning	.80	.78	.89	1.00	.86

Table 4.5 displays the intercorrelations between pairs of subscores within each test for both the college-bound and national groups. The disattenuated correlations among the Mathematics subscores and between the Usage/Mechanics and Rhetorical Skills subscores are fairly high. This result would indicate that, for example, if the true Usage/Mechanics subscore were known for an examinee, the examinee's true Rhetorical Skills subscore could be estimated with a high degree of accuracy. Note, however, that high *true* subscore intercorrelations does not rule out the possibility that the *observed* subscores may be useful, as distinct from the test score corresponding to the subscores, for some purposes. For any particular purpose the decision about the usefulness of subscores should be based on validity evidence specific to that purpose, rather than solely on true score correlations.

Table 4.5

Observed Correlations (Above Diagonal) and Disattenuated Correlations (Below Diagonal) Between Enhanced ACT Assessment Subscores

National

	Usage/Mechanics	Rhetorical Skills	English
Usage/Mechanics	1.00	.82	.96
Rhetorical Skills	.97	1.00	.94

	Pre-Algebra/ Elementary Algebra	Intermediate Algebra/ Coordinate Geometry	Plane Geometry/ Trigonometry	Mathematics
Pre-Algebra/ Elementary Algebra	1.00	.60	.66	.92
Intermediate Algebra/ Coordinate Geometry	.92	1.00	.54	.78
Plane Geometry/ Trigonometry	.95	.97	1.00	.82

	Reading Social Studies/ Sciences	Reading Arts/ Literature	Reading
Reading Social Studies/ Sciences	1.00	.64	.89
Reading Arts/Literature	.80	1.00	.91

College-Bound

	Usage/Mechanics	Rhetorical Skills	English
Usage/Mechanics	1.00	.82	.96
Rhetorical Skills	.97	1.00	.94

	Pre-Algebra/ Elementary Algebra	Intermediate Algebra/ Coordinate Geometry	Plane Geometry/ Trigonometry	Mathematics
Pre-Algebra/ Elementary Algebra	1.00	.62	.68	.93
Intermediate Algebra/ Coordinate Geometry	.92	1.00	.57	.80
Plane Geometry/ Trigonometry	.95	.97	1.00	.83

	Reading Social Studies/ Sciences	Reading Arts/ Literature	Reading
Reading Social Studies/ Sciences	1.00	.63	.89
Reading Arts/Literature	.78	1.00	.90

Comparisons of Groups Differing in Educational Background

Table 4.6 provides information about the relationship among scores, academic level, and educational aspiration. In this table, the means for the college-bound group are greater than the means for the national group for all 4 test scores, 7 subscores, and the Composite. This finding indicates that, as expected, ACT scores are related to educational aspirations: students with higher aspirations earn higher scores. Also as expected, ACT scores are related to grade level: examinees in higher grades earn higher test scores.

In addition to comparing development across grade levels, groups of examinees who were presumed to differ in their educational development for reasons other than their grade were also compared. One of the pieces of background data gathered from the examinees in the Academic Skills Study was the number of English and mathematics courses they had taken and/or planned to take during their high school careers. It was thought that for first-semester seniors, this information should provide fairly accurate indications of how many courses they had taken or were enrolled in. Of the 5,058 twelfth-grade examinees, 50 had invalid or missing responses to the item concerning English courses and 63 had invalid or missing responses to the item concerning mathematics courses. For each item, the remaining examinees were grouped into three categories: those who had taken and/or planned to take 3½ or more years of the subject, those who had taken and/or planned to take 2½ or 3 years of the subject, and those who had taken and/or planned to take 2 years or less of the subject.

Table 4.7A displays the means for those three groups for the item related to English courses. It can be seen that the largest differences are, not unexpectedly, between those who have taken and/or are planning to take at least 3½ years of English and those who have taken and/or are planning to take 2 years or less. Those who have taken and/or are planning to take between 2½ and 3 years performed more similarly to the 2-year-or-less group than to the 3½-or-more group. This pattern is also true for the mathematics item, as can be seen in Table 4.7B. Results are presented in more detail in Harris and Kolen (in press).

These findings--that students who have taken and/or plan to take more English and mathematics courses (and in the fall of grade 12 only one semester of courses at most is left to plan) have higher English and Mathematics test scores and subscores, are supportive of the enhanced ACT Assessment being a curriculum-based achievement test. However, it is also conceivable that able examinees take more English and mathematics classes and able examinees score higher on all types of English and mathematics tests.

Table 4.6

Scale Score Means and Standard Deviations

Test	No. Items	Grade 10	Grade 11	Grade 12
		National		
English	75	15.27(4.42)	16.24(4.82)	17.18(5.30)
Usage/Mechanics	(40)	7.29(2.94)	7.90(3.14)	8.50(3.42)
Rhetorical Skills	(35)	7.56(2.48)	8.08(2.72)	8.56(2.98)
Mathematics	60	15.68(2.86)	16.63(3.58)	17.44(4.45)
Pre-Alg./Elem. Alg.	(24)	7.45(2.65)	8.07(2.90)	8.56(3.31)
Int. Alg./Coor. Geo.	(18)	7.88(2.14)	8.25(2.38)	8.77(2.73)
Plane Geo./Trig.	(18)	7.76(2.36)	8.41(2.56)	8.73(2.85)
Reading	40	14.94(5.33)	16.17(5.89)	17.18(6.42)
Soc. Stu./Sci.	(20)	7.61(2.70)	8.14(2.98)	8.66(3.22)
Arts/Literature	(20)	7.03(4.02)	7.86(4.34)	8.40(4.52)
Science Reasoning	40	16.31(3.71)	16.91(3.93)	17.48(4.34)
Composite		15.67(3.51)	16.61(3.99)	17.45(4.54)
		College-Bound		
English	75	15.86(4.44)	16.91(4.84)	18.01(5.27)
Usage/Mechanics	(40)	7.65(2.97)	8.29(3.18)	9.01(3.40)
Rhetorical Skills	(35)	7.89(2.48)	8.47(2.71)	9.00(2.98)
Mathematics	60	16.00(2.92)	17.05(3.67)	18.00(4.56)
Pre-Alg/.Elem. Alg.	(24)	7.78(2.67)	8.44(2.90)	9.00(3.32)
Int. Alg./Coor. Geo.	(18)	8.01(2.11)	8.40(2.41)	9.00(2.78)
Plane Geo./Trig.	(18)	7.85(2.39)	8.59(2.60)	9.02(2.87)
Reading	40	15.59(5.38)	16.92(5.93)	18.01(6.47)
Soc. Stu./Sci.	(20)	7.84(2.77)	8.40(3.06)	8.98(3.30)
Arts/Literature	(20)	7.52(4.06)	8.45(4.35)	9.00(4.53)
Science Reasoning	40	16.72(3.74)	17.36(3.98)	17.99(4.41)
Composite		16.17(3.54)	17.18(4.02)	18.13(4.56)

Note: The values in this table differ slightly from those in Tables 3.5 and 3.6 because smoothed data were used in Tables 3.5 and 3.6 and unsmoothed data were used here.

Table 4.7

Means and Standard Deviations for Enhanced ACT Assessment Scores and Subscores: Grade 12 National Groups by Years of Coursework They Have Taken and/or Plan to Take

A. English

Test/ Subtest	No. Items	≤ 2 Yrs (N=126)*	2½-3 Yrs (N=221)*	≥ 3½ Yrs (N=4,661)*
English	75	13.26(3.28)	14.60(4.06)	17.46(5.32)
Usage/Mechanics	40	6.05(2.26)	7.02(2.70)	8.67(3.44)
Rhetorical Skills	35	6.56(1.86)	7.10(2.40)	8.71(3.00)

B. Mathematics

Test/ Subtest	No. Items	≤ 2 Yrs (N=758)*	2½-3 Yrs (N=1,642)*	≥ 3½ Yrs (N=2,595)*
Mathematics	60	14.61(2.37)	16.08(2.98)	19.35(4.91)
Pre-Alg./Elem. Alg.	24	6.32(2.12)	7.69(2.61)	9.96(3.41)
Int.Alg./Coor. Geo.	18	7.66(2.16)	8.08(2.19)	9.61(2.95)
Plane Geo./Trig.	18	7.43(2.13)	8.02(2.44)	9.67(2.99)

*N-counts are actual numbers of examinees; means and standard deviations are computed on weighted data.

Conclusion

This chapter presented reliability and validity evidence for the enhanced ACT Assessment tests of educational development. Reliability evidence, based on the Academic Skills Study, included KR20 reliability estimates for raw scores and estimated reliability coefficients for scale scores. Standard errors of measurement for scale scores were also reported. The standard errors of measurement are approximately 2 scale-score points throughout the scale score range for test scores, and approximately 1 for the Composite.

The validity evidence presented relates primarily to content validity. The limited statistical validity evidence currently available is in the form of relationships among test scores and differential test performance of groups with different educational backgrounds. Additional validity studies are currently being planned and will be implemented once the enhanced ACT Assessment Program has been administered operationally.

References

American College Testing Program. (1989). The enhanced ACT Assessment: Using concordance tables. Iowa City, IA: Author.

American College Testing Program. (1982). ACT/AAP and other admissions uses of the ACT Assessment Program by colleges and universities. Iowa City, IA: Author.

American College Testing Program. (1987). College student profiles: Norms for the ACT Assessment. Iowa City, IA: Author.

American College Testing Program. (1988). ACT Assessment program technical manual. Iowa City, IA: Author.

American Psychological Association. (1985). Standards for educational and psychological tests. Washington, DC: Author.

Cope, R. T., & Kolen, M. J. (1987, April). A study of methods for estimating distributions of test scores. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C.

Hanson, B. A. (1988, April). The use of smoothness assumptions in improving estimation of test score distributions. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Harris, D. J., & Kolen, M. J. (in press). Examining group differences from a developmental perspective. In R. L. Brennan (Ed.) Methodology used in scaling the ACT Assessment and P-ACT+. Iowa City, IA: The American College Testing Program.

Jarjoura, D., & Brennan, R. L. (1982). A variance components model for measurement procedures associated with a table of specifications. Applied Psychological Measurement, 6, 161-171.

Kolen, M. J. (1984). Effectiveness of analytic smoothing in equipercentile equating. Journal of Educational Statistics, 9, 25-44.

Kolen, M. J. (1988). Defining score scales in relation to measurement error. Journal of Educational Measurement, 25, 97-110.

Kolen, M. J., & Hanson, B. A. (in press). Scaling the ACT Assessment. In R. L. Brennan (Ed.) Methodology used in scaling the ACT Assessment and P-ACT+. Iowa City, IA: The American College Testing Program.

Lord, F. M. (1965). A strong true score theory, with applications. Psychometrika, 30, 239-270.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.) Educational Measurement (3rd edition). New York: American Council on Education, and Macmillan.

Sawyer, R. (in press). Research design and sampling for the Academic Skills Study. In R. L. Brennan (Ed.) Methodology used in scaling the ACT Assessment and P-ACT+. Iowa City, IA: The American College Testing Program.

Sawyer, R., Laing, J, & Houston, M. (1988). Accuracy of self-reported high school courses and grades of college-bound students. (Research Report No. 88-1). Iowa City: The American College Testing Program.

Wang, M. W., & Stanley, J. L. (1970). Differential weighting: A review of methods and empirical studies. Review of Educational Research, 40, 663-705.

APPENDIX A

ACT Interest Inventory

College planning usually involves career planning. Your scores on the ACT Interest Inventory will suggest educational programs and occupations you may want to think about—options you may not otherwise have considered.

Directions

Indicate how much you would like doing each of the activities listed below. Mark a response to an activity even if you are uncertain how you feel about it. Consider whether you would like or dislike an activity rather than whether you have the ability to do it.

For each question, choose one of the answers from the scale below and mark the corresponding letter on page 3 of your registration folder. Try to answer like or dislike to as many questions as possible.

I would dislike doing this activity ..... D
I am indifferent (don't care one way or the other) I
I would like doing this activity L

1. Learn about star formations
2. Sketch and draw pictures
3. Help someone make an important decision
4. Conduct a meeting
5. Count and sort money
6. Build a picture frame
7. Learn how the brain works
8. Compose or arrange music
9. Give first aid to an injured person
10. Develop new rules or policies
11. Take inventory in a store
12. Fix a toy
13. Explore a science museum
14. Make creative photographs
15. Show children how to play a game or sport
16. Work in a political campaign
17. Write payroll checks
18. Run a lawnmower
19. Attend the lecture of a well-known scientist
20. Write short stories
21. Work on a community improvement project
22. Present information before a group
23. Set up a bookkeeping system
24. Watch for forest fires
25. Study biology
26. Read about the writing style of modern authors
27. Help a newcomer meet people
28. Discuss a misleading advertisement with a salesperson
29. Prepare a budget for a club or group
30. Build furniture
31. Measure chemicals in a test tube
32. Prepare drawings to illustrate a magazine story
33. Take part in a small group discussion
34. Plan work for other people
35. Balance a checkbook
36. Learn to cut and polish gemstones
37. Read about a new surgical procedure
38. Write a movie script
39. Find out how others believe a problem can be solved
40. Explain legal rights to people
41. Sort, count, and store supplies
42. Repair damage to a tree after a storm
43. Study plant diseases
44. Select music to play for a local radio station
45. Help rescue someone in danger
46. Demonstrate a new product
47. Plan a monthly budget
48. Design a bird feeder
49. Read books or magazines about new scientific findings
50. Play jazz in a combo
51. Help settle an argument between friends
52. Campaign for a political office
53. Find errors in a financial account
54. Engrave lettering or designs on a trophy or plaque
55. Study chemistry
56. Draw cartoons
57. Give directions to visitors
58. Publicize a show or athletic event
59. Figure shipping costs for catalog orders
60. Operate a slide or movie projector
61. Use a microscope or other lab equipment
62. Design a metal sculpture
63. Help friends with their problems
64. Conduct business by phone
65. Make charts or graphs
66. Learn how to repair a typewriter
67. Read about the origin of the earth, sun, and stars
68. Play in a band
69. Teach people a new hobby
70. Interview workers about company complaints
71. Calculate the interest on a loan
72. Watch a technician repair a television
73. Observe and classify butterflies
74. Write reviews of Broadway plays
75. Help people during emergencies
76. Hire a person for a job
77. Keep expense account records
78. Prune plants and shrubs
79. Study the effects of vitamins on animals
80. Design a poster for an event
81. Entertain others by telling jokes or stories
82. Manage a small business
83. Look for errors in the draft of a report
84. Shelve books in a library
85. Learn how birds migrate
86. Play a musical instrument
87. Give a tour of an exhibit
88. Conduct a door-to-door opinion poll
89. Operate office machines
90. Inspect products for defects

BEST COPY AVAILABLE

APPENDIX B

For further information, refer to The Enhanced ACT Assessment: Using Concordance Tables (ACT, 1989).

Concordance Tables

ACT Composite Score

Composite score on current ACT Assessment	Composite score on Enhanced ACT Assessment	
	Concordant value	Estimated score interval
35	36	35-36
34	35	34-36
33	34	33-36
32	33	32-35
31	32	31-34
30	31	30-33
29	30	29-32
28	29	28-31
27	28	27-30
26	27	26-28
25	26	25-28
24	25	24-27
23	24	23-26
22	23	22-25
21	22	21-24
20	21	20-23
19	21	19-22
18	20	18-21
17	19	18-21
16	19	17-20
15	18	16-19
14	17	16-19
13	17	15-18
12	16	14-17
11	15	14-16
10	14	13-16
9	14	12-15
8	13	12-14
7	12	11-14
6	11	10-13
5	11	9-12
4	9	7-10
3	7	5-8
2	5	3-6
1	3	1-4

English

English Usage score on current ACT Assessment	English score on Enhanced ACT	
	Concordant value	Estimated score interval
33	36	33-36
32	35	32-36
31	34	31-36
30	34	30-35
29	33	29-34
28	32	27-32
27	31	26-31
26	30	25-30
25	29	23-29
24	28	22-28
23	27	21-27
22	25	20-25
21	24	19-24
20	22	18-23
19	21	17-22
18	20	16-21
17	19	15-20
16	18	15-20
15	17	14-19
14	16	13-18
13	15	13-18
12	15	12-17
11	14	11-17
10	13	11-16
9	13	11-16

Mathematics

Mathematics Usage score on current ACT Assessment	Mathematics score on Enhanced ACT	
	Concordant value	Estimated score interval
36	36	32-36
35	35	31-36
34	35	30-35
33	34	29-34
32	33	28-33
31	32	27-32
30	32	26-32
29	31	25-31
28	30	24-30
27	28	23-28
26	27	23-27
25	26	22-26
24	25	21-26
23	23	20-25
22	22	19-24
21	22	19-23
20	21	18-23
19	21	17-22
18	20	17-22
17	19	16-21
16	19	16-20
15	18	15-20
14	18	15-19
13	17	14-19
12	16	14-19
11	16	14-18
10	16	13-18
9	15	13-18